GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: February 17, 2004, 10:17:40 ; Search time 27 Seconds
(without alignments)
3301.791 Million cell updates/sec

Title: US-10-069-799-5
Perfect score: 927
Sequence: 1 MSNINVIKSNIQAGLNSTKS..........SSNALQPITQPTQGILAPSV 927

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 283308 seqs, 96168682 residues

Word size : 0

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : PIR_76:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	14	1.5	998	2	I41078	hemolysin - Escher
2	14	1.5	998	2	T00227	hemolysin A toxin
3	14	1.5	1023	1	LEECA	hemolysin A - Esch
4	14	1.5	1024	2	S10056	hemolysin A - Esch
5	14	1.5	1049	1	S51784	toxin III - Actino
6	14	1.5	1052	1	B49219	toxin III - Actino
7	11	1.2	956	1	B33389	toxin II - Actinob
8	11	1.2	956	1	A43834	toxin II - Actinob
9	11	1.2	1022	1	I39643	RTX-toxin I - Acti
10	10	1.1	953	1	B30169	leukotoxin A - Pas
11	10	1.1	955	1	A35254	leukotoxin A - Pas
12	10	1.1	1403	2	S77624	mannuronan C-5-epi
13	10	1.1	4936	2	AH2515	hypothetical prote
14	8	0.9	161	2	AG2664	ATP synthase B cha
15	8	0.9	161	2	F97446	hypothetical prote
16	8	0.9	208	2	S34238	leukotoxin A - Pas
17	8	0.9	234	2	AD0468	conserved hypothet
18	8	0.9	262	2	E70714	hypothetical prote
19	8	0.9	270	2	AC2987	conserved hypothet
20	8	0.9	270	2	E98296	hypothetical prote
21	8	0.9	275	2	JE0183	chitinase (EC 3.2.
22	8	0.9	314	2	JC7335	chitinase (EC 3.2.
23	8	0.9	317	2	AD3003	hypothetical prote
24	8	0.9	317	2	E98280	mpsA protein [impo
25	8	0.9	321	2	AB1054	probable membrane
26	8	0.9	325	2	C87464	lipoic acid syntha
27	8	0.9	353	2	E95914	probable secreted
28	8	0.9	368	2	AF1434	AA3-600 quinol oxi
29	8	0.9	368	2	AF1806	AA3-600 quinol oxi
30	8	0.9	380	2	AB3160	santhopine deamina
31	8	0.9	394	2	T19181	hypothetical prote
32	8	0.9	399	2	T19180	hypothetical prote
33	8	0.9	433	1	E64242	GTP-binding protei
34	8	0.9	443	2	AE0826	probable cadaverin
35	8	0.9	447	2	C84306	hypothetical prote
36	8	0.9	465	2	H95369	EglC ENDO-1,3-1,4-
37	8	0.9	469	2	B84372	DNA damage-inducib
38	8	0.9	474	2	S60902	CDP-ribitol pyroph
39	8	0.9	477	2	D84306	sodium- and chlori
40	8	0.9	505	2	G84079	sodium-dependent t
41	8	0.9	539	2	G95405	hypothetical prote
42	8	0.9	553	2	S77623	mannuronan C-5-epi
43	8	0.9	559	2	E84213	hypothetical prote
44	8	0.9	589	2	AD2263	hypothetical prote
45	8	0.9	643	2	T03518	hypothetical prote
46	8	0.9	773	2	S46011	probable pre-mRNA-
47	8	0.9	864	2	D70183	chemotaxis histidi
48	8	0.9	993	2	AE1905	outer membrane sec
49	8	0.9	997	2	S77625	mannuronan C-5-epi
50	8	0.9	997	2	I39739	mannuronan C5 epim
51	8	0.9	1055	1	A37205	leukotoxin A - Act
52	8	0.9	1112	2	H95964	probable outer mem
53	8	0.9	1208	2	C82779	hemolysin-type cal
54	8	0.9	1296	2	C82521	hemolysin-type cal
55	8	0.9	1417	2	AG2137	hypothetical prote
56	8	0.9	1705	2	S51672	adenylate cyclase
57	8	0.9	1736	2	F86178	hypothetical prote
58	8	0.9	1741	2	S74910	hemolysin - Synech
59	8	0.9	1839	2	S77626	mannuronan C-5-epi
60	8	0.9	1944	2	AH3098	rhizobiocin/RTX to
61	8	0.9	1990	2	A96188	probable phosphoes
62	8	0.9	3083	2	AH2493	hypothetical prote
63	7	0.8	17	2	I46151	histone H2AF - dog
64	7	0.8	17	2	I45826	histone H2AF - New
65	7	0.8	54	2	S22936	sex-determining pr
66	7	0.8	63	1	CKMTA	cecropin A precurs
67	7	0.8	63	2	D97714	hypothetical prote
68	7	0.8	79	2	B87640	hypothetical prote
69	7	0.8	79	2	S53337	CD47 glycoprotein
70	7	0.8	79	2	F97432	hypothetical prote
71	7	0.8	80	1	NTSR4T	neurotoxin TsIV pr
72	7	0.8	81	2	G87600	hypothetical prote
73	7	0.8	85	2	E83311	hypothetical prote
74	7	0.8	103	1	G64093	ribosomal protein
75	7	0.8	104	1	R5EC24	ribosomal protein
76	7	0.8	104	2	JC2277	ribosomal protein
77	7	0.8	104	2	F91150	50S ribosomal subu
78	7	0.8	104	2	B85996	50S ribosomal subu
79	7	0.8	104	2	AE0027	50S ribosomal prot
80	7	0.8	104	2	AD1007	50S ribosomal chai
81	7	0.8	105	2	C82058	ribosomal protein
82	7	0.8	110	2	T22476	hypothetical prote
83	7	0.8	110	2	F83745	hypothetical prote
84	7	0.8	112	2	H69537	conserved hypothet
85	7	0.8	112	2	AE2244	hypothetical prote
86	7	0.8	118	2	S51207	cruxhalorhodopsin-
87	7	0.8	118	2	S19120	pilin (pilS5 locus
88	7	0.8	122	2	AD0456	50S ribosomal prot
89	7	0.8	123	2	JE0093	histone H2A.F/Z va
90	7	0.8	123	2	H81235	50S ribosomal prot
91	7	0.8	125	2	S07392	histone H2A.F/Z -
92	7	0.8	125	2	A28343	sexual inducer gly
93	7	0.8	126	2	G84215	hypothetical prote
94	7	0.8	128	1	HSCH2F	histone H2A.F, emb
95	7	0.8	128	2	A35881	histone H2A.Z - hu
96	7	0.8	128	2	S03642	histone H2A.Z - bo
97	7	0.8	128	2	S03644	histone H2A.Z - ra
98	7	0.8	133	2	S60662	photosystem II pro
99	7	0.8	133	2	A84710	photosystem II rea
100	7	0.8	133	2	C97301	uncharacterized co
101	7	0.8	141	2	S08118	histone H2A.vD - f
102	7	0.8	141	2	S73848	hypothetical prote

103	7	0.8	147	2	D86389	hypothetical prote
104	7	0.8	149	2	T21037	hypothetical prote
105	7	0.8	151	2	H90256	hypothetical prote
106	7	0.8	154	2	A42418	auracyanin B1 - Ch
107	7	0.8	168	2	S57416	fimbrial protein M
108	7	0.8	168	2	A45943	vitelline membrane
109	7	0.8	172	2	E95958	hypothetical membr
110	7	0.8	181	2	S23604	hypothetical prote
111	7	0.8	182	2	S01150	H+-transporting tw
112	7	0.8	186	2	S35057	NADH2 dehydrogenas
113	7	0.8	187	2	S48361	transcription fact
114	7	0.8	189	2	AI0213	probable membrane
115	7	0.8	199	2	A72235	translation elonga
116	7	0.8	199	2	S75344	hypothetical prote
117	7	0.8	206	1	I64088	conserved hypothet
118	7	0.8	208	2	S01012	sexual inducer gly
119	7	0.8	213	2	B87149	probable carbonic
120	7	0.8	216	2	C75403	hypothetical prote
121	7	0.8	217	2	AF3050	two component resp
122	7	0.8	217	2	F98235	probable two-compo
123	7	0.8	218	2	S58769	brain-specific pro
124	7	0.8	218	2	H82758	hypothetical prote
125	7	0.8	219	2	G83137	riboflavin synthas
126	7	0.8	219	2	A70338	ABC transporter -
127	7	0.8	219	2	B95953	probable secreted
128	7	0.8	220	2	E84057	hypothetical prote
129	7	0.8	220	2	F69218	conserved hypothet
130	7	0.8	227	2	F83097	probable two-compo
131	7	0.8	229	1	G69066	probable 3,4-dihyd
132	7	0.8	232	2	S70355	phosphoribosylanth
133	7	0.8	237	2	C42606	orfB 3' to orf405
134	7	0.8	240	2	B45359	proopiomelanocorti
135	7	0.8	241	2	B75628	DNA-binding stress
136	7	0.8	247	2	JC2268	hemagglutinin - Ma
137	7	0.8	249	2	AI3156	short-chain dehydr
138	7	0.8	249	2	H98130	probable short-cha
139	7	0.8	251	2	AC0534	probable hydroxyac
140	7	0.8	254	2	S02066	transcription regu
141	7	0.8	254	2	E90969	hypothetical prote
142	7	0.8	254	2	D85742	hypothetical prote
143	7	0.8	256	2	S11878	chlorophyll a/b-bi
144	7	0.8	257	2	E82094	sodium-translocati
145	7	0.8	258	1	CTONPK	corticotropin / li
146	7	0.8	260	1	OTXL3	cytochrome-c oxida
147	7	0.8	261	2	T11540	cytochrome-c oxida
148	7	0.8	261	2	T11460	cytochrome-c oxida
149	7	0.8	261	2	T11826	cytochrome-c oxida
150	7	0.8	262	2	T34691	hypothetical prote
151	7	0.8	269	2	D70660	probable aroE - My
152	7	0.8	275	2	A22826	phytohemagglutinin
153	7	0.8	276	2	S51831	phytohemagglutinin
154	7	0.8	276	2	T48843	halorhodopsin [val
155	7	0.8	278	2	E83453	hypothetical prote
156	7	0.8	279	2	B71415	probable phosphoch
157	7	0.8	279	2	B87434	hypothetical prote
158	7	0.8	281	2	AG1369	conserved hypothet
159	7	0.8	282	2	T25037	hypothetical prote
160	7	0.8	284	2	S08385	nodO protein - Rhi
161	7	0.8	284	2	A43721	nodule formation p
162	7	0.8	291	2	AB1025	araC family regula
163	7	0.8	291	2	H72238	methicillin resist
164	7	0.8	294	2	AG0913	probable adenine-s
165	7	0.8	296	2	T26406	hypothetical prote
166	7	0.8	297	2	S75656	hypothetical prote
167	7	0.8	298	2	E85166	probable phosphoch
168	7	0.8	299	2	S76631	hypothetical prote
169	7	0.8	302	2	D82351	rarD protein VC019
170	7	0.8	307	2	S73632	Holliday junction
171	7	0.8	307	2	G64239	Holliday junction
172	7	0.8	311	2	G96596	hypothetical prote
173	7	0.8	311	2	D72693	hypothetical prote
174	7	0.8	312	2	T16001	hypothetical prote
175	7	0.8	314	2	F96527	protein F27J15.20
176	7	0.8	317	2	B83760	hypothetical prote
177	7	0.8	319	2	AB2121	cysteine synthase
178	7	0.8	319	2	F83402	binding protein co
179	7	0.8	319	2	S06367	ND1 intron 1 prote
180	7	0.8	320	2	S66176	ACC oxidase (clone
181	7	0.8	320	2	A40216	flavonol 4'-sulfot
182	7	0.8	320	2	AH2374	cysteine synthase
183	7	0.8	320	2	G96714	hypothetical prote
184	7	0.8	323	2	A48997	tumor surface anti
185	7	0.8	324	2	E83306	cysteine synthase
186	7	0.8	324	2	A70584	phosphate transpor
187	7	0.8	326	2	S54042	hypothetical prote
188	7	0.8	326	4	S61652	hypothetical prote
189	7	0.8	333	1	JS0590	endo-1,4-beta-xyla
190	7	0.8	333	2	I55593	connexin 37 - huma
191	7	0.8	334	2	A29561	prostatic spermine
192	7	0.8	335	2	T51106	dTDPglucose 4,6-de
193	7	0.8	338	2	H87510	polyprenyl synthet
194	7	0.8	338	2	B82128	conserved hypothet
195	7	0.8	338	2	G71128	hypothetical prote
196	7	0.8	339	2	AC3406	1-lactate permease
197	7	0.8	340	2	A55528	recombination prot
198	7	0.8	341	2	B86792	hypothetical prote
199	7	0.8	348	2	G83976	cytochrome caa3 ox
200	7	0.8	349	2	T16882	hypothetical prote
201	7	0.8	349	2	AC0419	probable transport
202	7	0.8	350	2	AC2375	hypothetical prote
203	7	0.8	351	2	E83872	hypothetical prote
204	7	0.8	351	2	D90264	biotin synthase (b
205	7	0.8	351	2	AC0302	conserved hypothet
206	7	0.8	353	2	JA0123	replicase - potato
207	7	0.8	355	2	T11600	NADH2 dehydrogenas
208	7	0.8	357	2	T19336	hypothetical prote
209	7	0.8	358	2	G70601	hypothetical prote
210	7	0.8	360	2	AD1929	hypothetical prote
211	7	0.8	362	2	F97847	recA protein [impo
212	7	0.8	363	2	S51104	outer membrane por
213	7	0.8	364	2	AE2611	conserved hypothet
214	7	0.8	364	2	C97393	hypothetical prote
215	7	0.8	366	2	B81055	chorismate synthas
216	7	0.8	366	2	A81822	chorismate synthas
217	7	0.8	367	2	C82391	N-ethylmaleimide r
218	7	0.8	367	2	T12790	N-acetylmuramoyl-L
219	7	0.8	371	2	T51695	cell division prot
220	7	0.8	374	2	E82690	chorismate synthas
221	7	0.8	375	2	E75542	probable endogluca
222	7	0.8	375	2	JQ0846	DNA-binding protei
223	7	0.8	376	2	S73941	oligopeptide trans
224	7	0.8	379	2	G75428	conserved hypothet
225	7	0.8	380	2	T42755	tyrosylprotein sul
226	7	0.8	382	2	E64686	myosin-like protei
227	7	0.8	384	2	T38544	probable exopolyph
228	7	0.8	385	2	D95350	probable transmemb
229	7	0.8	386	2	D71851	tetracycline resis
230	7	0.8	386	2	E64665	tetracycline resis
231	7	0.8	387	2	E95933	probable calcium-b
232	7	0.8	388	2	AC3216	isomerase/lactoniz
233	7	0.8	388	2	AG2723	phage-related inte
234	7	0.8	389	2	A60109	major outer membra
235	7	0.8	393	2	AF0365	probable aminotran
236	7	0.8	395	2	H90441	hypothetical prote
237	7	0.8	397	2	F83385	probable MFS trans
238	7	0.8	400	2	C82855	aromatic-amino-aci
239	7	0.8	401	2	B84060	multidrug resistan
240	7	0.8	404	2	G69318	threonine synthase
241	7	0.8	406	1	JNEBPT	phosphoglycerate t
242	7	0.8	406	2	C43664	tryptophan synthas
243	7	0.8	407	2	B70185	probable hydroxyme
244	7	0.8	407	2	D86782	peptidoglycan bios
245	7	0.8	410	2	F87688	tryptophan synthas
246	7	0.8	410	2	G85991	hypothetical prote
247	7	0.8	411	2	H95851	probable alcohol d
248	7	0.8	414	1	JQ1016	glycine hydroxymet

249	7	0.8	414	2	H81383	glycine hydroxymet
250	7	0.8	414	2	E83489	hypothetical prote
251	7	0.8	415	2	AI2087	hypothetical prote
252	7	0.8	426	2	E87385	chromate transport
253	7	0.8	429	2	C97023	homoserine dehydro
254	7	0.8	430	2	C97237	histidyl-tRNA synt
255	7	0.8	431	2	G86277	F14L17.11 protein
256	7	0.8	439	2	T01270	hypothetical prote
257	7	0.8	441	2	G75306	drug transport pro
258	7	0.8	442	2	H70960	hypothetical prote
259	7	0.8	443	2	AF1901	hypothetical prote
260	7	0.8	444	2	E83033	probable MFS trans
261	7	0.8	449	2	E72383	dihydrolipoamide d
262	7	0.8	449	2	G64349	phosphomannomutase
263	7	0.8	449	2	B86763	hypothetical prote
264	7	0.8	450	2	C86262	F13K23.12 protein
265	7	0.8	452	2	H90249	pyruvate kinase (p
266	7	0.8	452	2	B59096	hypothetical prote
267	7	0.8	452	2	F75217	hypothetical prote
268	7	0.8	454	2	E75116	4-aminobutyrate am
269	7	0.8	454	2	A71016	probable 4-aminobu
270	7	0.8	455	2	G71860	biotin carboxylase
271	7	0.8	458	2	B64566	biotin carboxylase
272	7	0.8	460	2	B83552	probable two-compo
273	7	0.8	460	2	I57546	Rabin3 - rat
274	7	0.8	462	2	A83893	two-component sens
275	7	0.8	463	2	AB0807	phosphoglycerate t
276	7	0.8	463	2	T51313	hypothetical prote
277	7	0.8	465	2	AB2995	endo-1,3-1,4-beta-
278	7	0.8	466	2	T45880	hypothetical prote
279	7	0.8	467	2	D84938	H+-transporting tw
280	7	0.8	471	2	T34956	probable UDP-N-ace
281	7	0.8	473	2	G90401	flagella-related p
282	7	0.8	473	2	AE3584	exopolyphosphatase
283	7	0.8	474	2	AB0002	probable membrane
284	7	0.8	475	2	S48132	metalloproteinase
285	7	0.8	477	2	D82587	cationic amino aci
286	7	0.8	477	4	I41144	aspartate ammonia-
287	7	0.8	478	1	UFECDW	aspartate ammonia-
288	7	0.8	479	2	A38307	metalloproteinase
289	7	0.8	482	2	T08256	cytochrome d ubiqu
290	7	0.8	482	2	H98288	endo-1,3-1,4-beta-
291	7	0.8	483	2	S37055	catalase (EC 1.11.
292	7	0.8	483	2	G87412	outer membrane pro
293	7	0.8	484	2	C82171	cardiolipin syntha
294	7	0.8	485	2	T25199	hypothetical prote
295	7	0.8	486	1	HYSE15	serralysin (EC 3.4
296	7	0.8	488	1	H64313	corrinoid/iron-sul
297	7	0.8	492	2	A55589	lysine-tRNA ligase
298	7	0.8	493	2	H91268	aspartate ammonia-
299	7	0.8	493	2	F86109	aspartate ammonia-
300	7	0.8	500	2	B83910	hypothetical prote
301	7	0.8	504	2	S12164	serralysin (EC 3.4
302	7	0.8	505	2	H85361	leucyl aminopeptid
303	7	0.8	505	2	T04143	CLB1 protein - tom
304	7	0.8	506	2	T22876	hypothetical prote
305	7	0.8	510	2	T50526	CaLB protein - Ara
306	7	0.8	512	2	H70707	probable ggtA prot
307	7	0.8	525	2	A87468	malate synthase (i
308	7	0.8	530	2	T21171	hypothetical prote
309	7	0.8	532	1	QFPGM	neurofilament trip
310	7	0.8	534	2	F72682	hypothetical prote
311	7	0.8	547	2	A48442	membrane transport
312	7	0.8	557	1	S34412	transcription fact
313	7	0.8	557	1	S15342	transcription fact
314	7	0.8	557	2	T47128	heat shock protein
315	7	0.8	558	1	A39633	transcription fact
316	7	0.8	558	2	E72220	conserved hypothet
317	7	0.8	559	2	S33724	transcription fact
318	7	0.8	559	2	D95913	probable secreted
319	7	0.8	575	2	T48224	probable homeodoma
320	7	0.8	576	2	A70900	hypothetical glyci
321	7	0.8	583	2	AI3098	hypothetical prote
322	7	0.8	585	1	F64159	hypothetical prote
323	7	0.8	587	2	E82155	sigma-54 dependent
324	7	0.8	589	2	T29895	hypothetical prote
325	7	0.8	589	2	B97806	hypothetical prote
326	7	0.8	594	2	C81011	conserved hypothet
327	7	0.8	594	2	H81954	probable periplasm
328	7	0.8	595	2	T26843	hypothetical prote
329	7	0.8	600	2	B83875	two-component sens
330	7	0.8	606	2	H98187	prsD protein (U891
331	7	0.8	608	2	B82938	zinc metalloprotei
332	7	0.8	609	2	B81385	probable ATP /GTP
333	7	0.8	616	2	AG2957	hypothetical prote
334	7	0.8	616	2	G98325	probable c4-dicarb
335	7	0.8	617	2	C75407	excinuclease ABC c
336	7	0.8	620	2	F72395	hypothetical prote
337	7	0.8	632	2	AG1366	phosphotransferase
338	7	0.8	632	2	AH1735	phosphotransferase
339	7	0.8	635	2	F90551	lipoprotein (impor
340	7	0.8	637	2	T50951	hypothetical prote
341	7	0.8	639	2	A97450	1-deoxy-D-xylulose
342	7	0.8	639	2	AC2668	1-deoxy-D-xylulose
343	7	0.8	643	2	AD3439	1-deoxyxylulose-5-
344	7	0.8	646	2	S01351	type III site-spec
345	7	0.8	646	2	T23039	hypothetical prote
346	7	0.8	654	2	H86717	hypothetical prote
347	7	0.8	655	2	T34705	hypothetical prote
348	7	0.8	658	2	H81676	dnaK protein TC067
349	7	0.8	658	2	D96656	hypothetical prote
350	7	0.8	660	2	A40158	dnaK-type molecula
351	7	0.8	660	2	B71521	dnaK-type molecula
352	7	0.8	662	2	T32821	hypothetical prote
353	7	0.8	670	2	A25001	catalase (EC 1.11.
354	7	0.8	671	2	E82355	ATP-dependent DNA
355	7	0.8	676	2	AI2058	hypothetical prote
356	7	0.8	685	2	G97351	K+-transporting AT
357	7	0.8	685	2	T46843	H+/K+-exchanging A
358	7	0.8	689	2	AC1408	transcription anti
359	7	0.8	689	2	AC1784	transcription anti
360	7	0.8	698	2	S49206	G1 cyclin CLN1 - y
361	7	0.8	703	2	A49994	sulfate transporte
362	7	0.8	711	2	A90023	DNA topoisomerase
363	7	0.8	726	2	B72665	hypothetical prote
364	7	0.8	735	2	S76425	hypothetical prote
365	7	0.8	745	2	G01025	serine/threonine p
366	7	0.8	746	2	B70359	transcription regu
367	7	0.8	756	2	C87432	hypothetical prote
368	7	0.8	757	2	F83516	hypothetical prote
369	7	0.8	764	2	T05409	hypothetical prote
370	7	0.8	767	2	C97763	hypothetical prote
371	7	0.8	781	1	VCPV19	coat protein VP1 -
372	7	0.8	784	2	PN0009	neurofilament trip
373	7	0.8	790	2	AE2203	hypothetical prote
374	7	0.8	792	2	S03232	hypothetical prote
375	7	0.8	806	2	G95362	probable (imported
376	7	0.8	808	2	H64474	hypothetical prote
377	7	0.8	815	2	H69775	hypothetical prote
378	7	0.8	817	1	TGHUM1	protein-glutamine
379	7	0.8	827	1	S10639	fructose phosphotr
380	7	0.8	843	2	T34618	NADH2 dehydrogenas
381	7	0.8	847	2	T19544	hypothetical prote
382	7	0.8	850	2	G70332	conserved hypothet
383	7	0.8	855	2	A71223	hypothetical prote
384	7	0.8	855	2	B75191	probable ATP-depen
385	7	0.8	899	2	D96594	unknown protein, 7
386	7	0.8	900	2	T41607	probable vacuolar
387	7	0.8	900	2	AC1852	hypothetical prote
388	7	0.8	906	2	S38723	glutamate receptor
389	7	0.8	962	2	T04124	receptor-like prot
390	7	0.8	993	2	C55226	cylM protein - Ent
391	7	0.8	1014	2	H81894	probable cell-divi
392	7	0.8	1014	2	E81125	cell division prot
393	7	0.8	1030	2	S57380	probable membrane
394	7	0.8	1034	2	D65119	acriflavin resista

395	7	0.8	1038	2	JC5497	claustrin - chicke
396	7	0.8	1072	2	G95851	probable hemolysin
397	7	0.8	1083	2	A38919	hypothetical prote
398	7	0.8	1109	2	A56143	surface-array prot
399	7	0.8	1131	2	D75429	sensory box sensor
400	7	0.8	1138	2	A48944	parasporal crystal
401	7	0.8	1146	2	S59376	hypothetical prote
402	7	0.8	1170	2	AG3595	transcription-repa
403	7	0.8	1202	2	S55553	LAR-interacting pr
404	7	0.8	1204	2	T18812	hypothetical prote
405	7	0.8	1208	2	T42574	DNA-binding protei
406	7	0.8	1209	1	DNBEC4	DNA-binding protei
407	7	0.8	1237	2	E86457	probable RNA helic
408	7	0.8	1287	2	B53739	vacuolating cytoto
409	7	0.8	1288	2	E71884	vacuolating cytoto
410	7	0.8	1289	2	E90098	RNA polymerase III
411	7	0.8	1290	2	G64630	vacuolating cytoto
412	7	0.8	1291	1	A48940	bontoxilysin (EC 3
413	7	0.8	1291	2	S44983	vacuolating cytoto
414	7	0.8	1305	2	D82923	DNA-directed RNA p
415	7	0.8	1428	2	AC2224	hypothetical prote
416	7	0.8	1525	2	T14961	hypothetical prote
417	7	0.8	1609	2	E87243	probable cation tr
418	7	0.8	1632	2	C70752	probable ctpI prot
419	7	0.8	1636	2	B82736	hemolysin-type cal
420	7	0.8	1774	2	T17421	polyketide synthas
421	7	0.8	1830	2	E82909	conserved hypothet
422	7	0.8	2064	2	G82562	bacteriocin XF2407
423	7	0.8	2171	2	E86342	hypothetical prote
424	7	0.8	2364	2	A56577	microtubule-associ
425	7	0.8	2427	2	T16613	hypothetical prote
426	7	0.8	2464	1	QRMSP1	microtubule-associ
427	7	0.8	2761	2	T21064	hypothetical prote
428	7	0.8	2894	2	C64474	hypothetical prote
429	7	0.8	3135	2	A48584	transmission block
430	7	0.8	3430	1	GNWVWV	genome polyprotein
431	7	0.8	3433	1	GNWVKV	genome polyprotein
432	7	0.8	3436	2	S55659	tegument protein 6
433	7	0.8	3678	2	S28916	dystrophin - mouse
434	6	0.6	12	2	A29169	phospholipase A2 (
435	6	0.6	17	2	A29834	trp leader peptide
436	6	0.6	23	2	B54226	light-harvesting p
437	6	0.6	26	2	G37396	pollen allergen An
438	6	0.6	26	4	I59183	hypothetical serin
439	6	0.6	31	2	A58793	relaxin chain B -
440	6	0.6	31	2	A34448	myosin heavy chain
441	6	0.6	31	2	S65418	pyruvate synthase
442	6	0.6	35	2	D97553	hypothetical prote
443	6	0.6	36	2	C84175	hypothetical prote
444	6	0.6	37	2	S11202	ribosomal protein
445	6	0.6	38	2	S07140	ribosomal protein
446	6	0.6	42	2	S21027	M-factor m1 - fiss
447	6	0.6	45	2	A38666	lactogen I variant
448	6	0.6	51	2	S71030	hypothetical prote
449	6	0.6	52	2	B90673	probable invertase
450	6	0.6	52	2	F85523	probable invertase
451	6	0.6	54	2	AC0896	conserved hypothet
452	6	0.6	55	2	D97080	probable transcrip
453	6	0.6	57	2	F97843	hypothetical prote
454	6	0.6	58	2	AC0327	hypothetical prote
455	6	0.6	59	2	S72821	transcription init
456	6	0.6	59	2	D97219	hypothetical prote
457	6	0.6	59	2	D98151	hypothetical prote
458	6	0.6	60	1	LBRFA8	light-harvesting p
459	6	0.6	60	2	F90769	hypothetical prote
460	6	0.6	60	2	A64006	hypothetical prote
461	6	0.6	60	2	F97702	hypothetical prote
462	6	0.6	61	2	F95014	hypothetical prote
463	6	0.6	61	2	H97887	hypothetical prote
464	6	0.6	64	2	H81811	hypothetical prote
465	6	0.6	64	2	H97099	small acid-soluble
466	6	0.6	65	1	B25234	small acid-soluble
467	6	0.6	65	2	T28606	hypothetical prote

468	6	0.6	65	2	S19568	parsin, ovary-matu
469	6	0.6	65	2	G72172	H8R protein - vari
470	6	0.6	66	2	D82805	hypothetical prote
471	6	0.6	67	2	B70214	hypothetical prote
472	6	0.6	68	2	D85716	probable holin pro
473	6	0.6	68	2	D90766	probable holin pro
474	6	0.6	68	2	T17853	hypothetical prote
475	6	0.6	69	2	B46238	hypothetical prote
476	6	0.6	70	2	C95284	hypothetical prote
477	6	0.6	71	1	H70799	integrase-related
478	6	0.6	71	2	B90357	hypothetical prote
479	6	0.6	71	2	AG1985	hypothetical prote
480	6	0.6	74	2	H90657	hypothetical prote
481	6	0.6	74	2	H85508	hypothetical prote
482	6	0.6	74	2	AD0231	hypothetical prote
483	6	0.6	75	2	AH1550	hypothetical prote
484	6	0.6	75	2	AI2226	hypothetical prote
485	6	0.6	76	1	C69331	hypothetical prote
486	6	0.6	76	2	T36168	hypothetical prote
487	6	0.6	76	2	D83916	hypothetical prote
488	6	0.6	76	2	D82787	hypothetical prote
489	6	0.6	78	2	E85682	unknown protein en
490	6	0.6	78	2	AB1162	flagellar switch p
491	6	0.6	78	2	AB1521	weakly flagellar s
492	6	0.6	79	2	A86517	hypothetical prote
493	6	0.6	79	2	F83880	acyl-carrier prote
494	6	0.6	79	2	B72106	hypothetical prote
495	6	0.6	79	2	A83944	hypothetical prote
496	6	0.6	80	2	C49050	T-cell surface gly
497	6	0.6	80	2	T08876	phosphoribosylpyro
498	6	0.6	80	2	AB3302	hypothetical prote
499	6	0.6	81	1	JN0462	adenoregulin precu
500	6	0.6	81	2	A75384	hypothetical prote
501	6	0.6	81	2	D81565	hypothetical prote
502	6	0.6	83	2	A64463	hypothetical prote
503	6	0.6	83	2	A45271	invasin secretion
504	6	0.6	83	2	AI0303	hypothetical prote
505	6	0.6	84	1	A26873	small acid-soluble
506	6	0.6	84	1	QIBPS2	uracil-DNA glycosy
507	6	0.6	84	2	A64172	exodeoxyribonuclea
508	6	0.6	84	2	F90941	hypothetical prote
509	6	0.6	84	2	B85790	hypothetical prote
510	6	0.6	84	2	A64940	hypothetical prote
511	6	0.6	85	2	G83302	conserved hypothet
512	6	0.6	85	2	AD2229	hypothetical prote
513	6	0.6	86	2	T17859	hypothetical prote
514	6	0.6	86	2	S73048	hypothetical prote
515	6	0.6	86	2	S20471	class V zygote-spe
516	6	0.6	87	2	AH1828	hypothetical prote
517	6	0.6	88	2	A82920	hypothetical prote
518	6	0.6	89	2	D72581	hypothetical prote
519	6	0.6	92	2	T36228	hypothetical prote
520	6	0.6	93	2	T01876	hypothetical prote
521	6	0.6	94	2	S24795	ubiquinol-cytochro
522	6	0.6	94	2	T06417	pathogenesis-relat
523	6	0.6	94	2	S72613	chaperonin groES -
524	6	0.6	94	2	T30377	hypothetical prote
525	6	0.6	94	2	T03285	anther-specific pr
526	6	0.6	94	2	F95917	hypothetical prote
527	6	0.6	95	2	E69422	hypothetical prote
528	6	0.6	95	2	T03697	SAR8.2a protein, T
529	6	0.6	95	2	T03699	SAR8.2c protein, T
530	6	0.6	95	2	T03700	SAR8.2d protein, T
531	6	0.6	95	2	T02969	hypothetical prote
532	6	0.6	95	2	A83880	acyl-carrier prote
533	6	0.6	95	2	AI1024	probable GerE fami
534	6	0.6	95	2	C98255	hypothetical prote
535	6	0.6	96	2	G81019	chaperonin, 10 kDa
536	6	0.6	96	2	A89891	conserved hypothet
537	6	0.6	96	2	C69225	conserved hypothet
538	6	0.6	96	2	E83815	hypothetical prote
539	6	0.6	96	2	B82321	conserved hypothet
540	6	0.6	96	2	S10069	regulatory protein

541	6	0.6	96	2	A64626 hypothetical prote
542	6	0.6	97	2	E87330 S4 domain protein
543	6	0.6	98	1	T17351 NADH2 dehydrogenas
544	6	0.6	98	2	T17136 NADH2 dehydrogenas
545	6	0.6	98	2	T17143 NADH2 dehydrogenas
546	6	0.6	98	2	T17174 NADH2 dehydrogenas
547	6	0.6	98	2	T17171 NADH2 dehydrogenas
548	6	0.6	98	2	T17177 NADH2 dehydrogenas
549	6	0.6	98	2	T17183 NADH2 dehydrogenas
550	6	0.6	98	2	T17348 NADH2 dehydrogenas
551	6	0.6	98	2	T17360 NADH2 dehydrogenas
552	6	0.6	98	2	T17363 NADH2 dehydrogenas
553	6	0.6	98	2	T17147 NADH2 dehydrogenas
554	6	0.6	98	2	T17154 NADH2 dehydrogenas
555	6	0.6	98	2	S62433 zinc finger protei
556	6	0.6	98	2	C97749 hypothetical prote
557	6	0.6	99	2	T37184 hypothetical prote
558	6	0.6	99	2	T44977 oligopeptide ABC t
559	6	0.6	100	2	T30673 hypothetical prote
560	6	0.6	100	2	C64575 hypothetical prote
561	6	0.6	101	2	E82849 conserved hypothet
562	6	0.6	101	2	T45325 hypothetical prote
563	6	0.6	101	2	D97386 hypothetical prote
564	6	0.6	101	2	AD2604 conserved hypothet
565	6	0.6	101	2	S04714 hypothetical prote
566	6	0.6	101	2	S30493 Spl protein - mous
567	6	0.6	101	2	B96904 hypothetical prote
568	6	0.6	101	2	A97625 hypothetical prote
569	6	0.6	102	1	QPTK acylphosphatase (E
570	6	0.6	102	1	QPCH acylphosphatase (E
571	6	0.6	102	2	AF3395 NADH2 dehydrogenas
572	6	0.6	102	2	A45675 acylphosphatase (E
573	6	0.6	102	2	C87282 conserved hypothet
574	6	0.6	102	2	T25332 hypothetical prote
575	6	0.6	102	2	T36095 hypothetical prote
576	6	0.6	103	1	R5BS2B ribosomal protein
577	6	0.6	103	2	T11985 ribosomal protein
578	6	0.6	103	2	AB3348 LSU ribosomal prot
579	6	0.6	103	2	S61081 M protein precurso
580	6	0.6	103	2	C72307 conserved hypothet
581	6	0.6	103	2	S26067 hypothetical prote
582	6	0.6	103	2	B47048 plasmid stabilizat
583	6	0.6	103	2	G86911 hypothetical prote
584	6	0.6	103	2	H90234 hypothetical prote
585	6	0.6	103	2	C26424 stylar glycoprotei
586	6	0.6	104	2	S53942 probable membrane
587	6	0.6	104	2	E69805 hypothetical prote
588	6	0.6	104	2	H90411 hypothetical prote
589	6	0.6	105	1	S40199 ribosomal protein
590	6	0.6	105	2	G95911 probable ferredoxi
591	6	0.6	105	2	F72678 hypothetical prote
592	6	0.6	105	2	A82985 hypothetical prote
593	6	0.6	105	2	T37176 hypothetical prote
594	6	0.6	105	2	G72705 hypothetical prote
595	6	0.6	106	2	C64111 probable translati
596	6	0.6	106	2	H72547 probable periplasm
597	6	0.6	106	2	D72692 hypothetical prote
598	6	0.6	106	2	S72969 hypothetical 11.9K
599	6	0.6	106	2	S72815 hypothetical prote
600	6	0.6	106	2	T06685 hypothetical prote
601	6	0.6	106	2	PN0468 hypothetical prote
602	6	0.6	107	2	S75164 hypothetical prote
603	6	0.6	107	2	JQ2035 hypothetical 11.7K
604	6	0.6	108	2	E75355 hypothetical prote
605	6	0.6	108	2	E64962 hypothetical prote
606	6	0.6	109	2	S61255 genome polyprotein
607	6	0.6	109	2	T02762 anther specific pr
608	6	0.6	110	2	PH0211 probable transamin
609	6	0.6	110	2	D97016 phosphoribosyl-ATP
610	6	0.6	110	2	D83170 hypothetical prote
611	6	0.6	110	2	E72423 hypothetical prote
612	6	0.6	110	2	E72663 hypothetical prote
613	6	0.6	111	2	T06418 pathogenesis-relat

614	6	0.6	111	2	H90263 transcriptional re
615	6	0.6	111	2	T03701 SAR8.2e protein, T
616	6	0.6	112	2	AB2722 integration host f
617	6	0.6	112	2	B69468 nitrogen regulator
618	6	0.6	112	2	S30269 protein hdeB precu
619	6	0.6	112	2	E91177 hypothetical prote
620	6	0.6	112	2	F86023 hypothetical prote
621	6	0.6	112	2	E85686 unknown protein en
622	6	0.6	113	2	T30494 hypothetical prote
623	6	0.6	113	2	S76866 hypothetical prote
624	6	0.6	113	2	T36789 hypothetical prote
625	6	0.6	114	1	RIIDS1 somatostatin-14 pr
626	6	0.6	114	2	I50798 preprosomatostatin
627	6	0.6	114	2	C72759 hypothetical prote
628	6	0.6	115	2	A23329 Ly-5-8 glycoprotei
629	6	0.6	115	2	S06809 keratin - chicken
630	6	0.6	115	2	AE3313 protein secretion
631	6	0.6	115	2	T30042 hypothetical prote
632	6	0.6	116	1	R5YM18 ribosomal protein
633	6	0.6	116	2	G97503 himA protein [impo
634	6	0.6	116	2	E72551 hypothetical prote
635	6	0.6	116	2	D81945 probable pilus bio
636	6	0.6	116	2	F85957 unknown protein en
637	6	0.6	116	2	E85605 hypothetical prote
638	6	0.6	116	2	B90796 hypothetical prote
639	6	0.6	116	2	D91112 hypothetical prote
640	6	0.6	116	2	H90190 probable transcrip
641	6	0.6	116	2	T08906 hypothetical prote
642	6	0.6	117	2	D81442 50S ribosomal prot
643	6	0.6	117	2	S19123 hypothetical prote
644	6	0.6	117	2	B71456 hypothetical prote
645	6	0.6	117	2	T42313 hypothetical prote
646	6	0.6	117	2	T12498 hypothetical prote
647	6	0.6	118	2	C70382 ribosomal protein
648	6	0.6	118	2	S01253 hypothetical prote
649	6	0.6	118	2	I64028 hypothetical prote
650	6	0.6	119	2	F75110 hypothetical prote
651	6	0.6	119	2	B75331 conserved hypothet
652	6	0.6	120	2	D75324 hypothetical prote
653	6	0.6	121	1	R5EC7 ribosomal protein
654	6	0.6	121	1	R5EB12 ribosomal protein
655	6	0.6	121	2	E91242 50S ribosomal subu
656	6	0.6	121	2	E82336 ribosomal protein
657	6	0.6	121	2	B82946 ribosomal protein
658	6	0.6	121	2	B86090 50S ribosomal subu
659	6	0.6	121	2	AF0933 50S ribosomal chai
660	6	0.6	121	2	E71345 conserved hypothet
661	6	0.6	122	1	R5HG12 ribosomal protein
662	6	0.6	122	2	A24753 ribosomal protein
663	6	0.6	122	2	F82533 50S ribosomal prot
664	6	0.6	122	2	S49638 hypothetical prote
665	6	0.6	122	2	D87555 hypothetical prote
666	6	0.6	122	2	B97081 iron-dependent tra
667	6	0.6	123	1	VHWWHE structural protein
668	6	0.6	123	1	C44212 structural protein
669	6	0.6	123	2	F64083 ribosomal protein
670	6	0.6	123	2	G97286 ribosomal protein
671	6	0.6	123	2	G83980 transcription regu
672	6	0.6	123	2	D90488 conserved hypothet
673	6	0.6	123	2	S75382 hypothetical prote
674	6	0.6	123	2	F82716 hypothetical prote
675	6	0.6	124	2	I40348 ribosomal protein
676	6	0.6	124	2	I40350 ribosomal protein
677	6	0.6	124	2	AF3345 LSU ribosomal prot
678	6	0.6	124	2	B70472 flagellar protein
679	6	0.6	124	2	B96960 holo-acyl-carrier
680	6	0.6	124	2	T13559 hypothetical prote
681	6	0.6	125	1	CCAGA6 cytochrome c556 -
682	6	0.6	125	2	H71723 ribosomal protein
683	6	0.6	125	2	AC2817 50S ribosomal prot
684	6	0.6	125	2	D97595 ribosomal protein
685	6	0.6	125	2	B97022 general secretion
686	6	0.6	125	2	S23214 hypothetical prote

687	6	0.6	126	1	R5DV7	ribosomal protein
688	6	0.6	126	2	F69787	hypothetical prote
689	6	0.6	126	2	D69293	conserved hypothet
690	6	0.6	127	2	S53854	ribosomal protein
691	6	0.6	127	2	H87310	ribosomal protein
692	6	0.6	127	2	E86158	60S ribosomal prot
693	6	0.6	127	2	T47929	hypothetical prote
694	6	0.6	127	2	AG2667	conserved hypothet
695	6	0.6	127	2	D97449	hypothetical prote
696	6	0.6	127	2	PC2036	microfibril-associ
697	6	0.6	127	2	AG1425	hypothetical secre
698	6	0.6	128	2	D71906	diacylglycerol kin
699	6	0.6	128	2	D64607	diacylglycerol kin
700	6	0.6	128	2	S77671	streptokinase A (E
701	6	0.6	128	2	S77679	streptokinase A (E
702	6	0.6	128	2	S77688	streptokinase A (E
703	6	0.6	128	2	AE3399	integration host f
704	6	0.6	128	2	E70466	ribosomal protein
705	6	0.6	128	2	S13069	ribosomal protein
706	6	0.6	128	2	JQ1002	keratin, claw - ch
707	6	0.6	128	2	AC0111	glycine cleavage s
708	6	0.6	129	2	S72292	ribosomal protein
709	6	0.6	129	2	D64444	hypothetical prote
710	6	0.6	129	2	H82735	hypothetical prote
711	6	0.6	129	2	F90349	conserved hypothet
712	6	0.6	129	2	T20172	hypothetical prote
713	6	0.6	130	2	AG2468	50S ribosomal prot
714	6	0.6	130	2	AE1511	hypothetical prote
715	6	0.6	130	2	H84943	hypothetical prote
716	6	0.6	130	2	D84507	hypothetical prote
717	6	0.6	130	2	AC1799	hypothetical secre
718	6	0.6	131	2	T07390	14-3-3 protein tft
719	6	0.6	131	2	H97780	hypothetical prote
720	6	0.6	131	2	F90153	hypothetical prote
721	6	0.6	131	2	D82222	hypothetical prote
722	6	0.6	132	2	T18337	icmD protein - Leg
723	6	0.6	132	2	S69753	hypothetical prote
724	6	0.6	132	2	C83298	hypothetical prote
725	6	0.6	132	2	T23596	hypothetical prote
726	6	0.6	133	1	GPBPP4	gop protein - sate
727	6	0.6	133	2	D87690	hypothetical prote
728	6	0.6	133	2	A70986	hypothetical prote
729	6	0.6	134	2	B90284	hemoglobin alpha-2
730	6	0.6	134	2	S66694	histone H2A.HTA3 -
731	6	0.6	134	2	B71359	conserved hypothet
732	6	0.6	134	2	F64015	hypothetical prote
733	6	0.6	134	2	S77129	hypothetical prote
734	6	0.6	134	2	B83804	small multidrug ex
735	6	0.6	135	2	AF0865	conserved hypothet
736	6	0.6	135	2	T28703	hypothetical prote
737	6	0.6	135	2	H87410	hypothetical prote
738	6	0.6	135	2	I64147	lactoylglutathione
739	6	0.6	135	2	D72612	hypothetical prote
740	6	0.6	136	2	T10795	acyl carrier prote
741	6	0.6	136	2	C81436	biopolymer transpo
742	6	0.6	136	2	A65083	hypothetical prote
743	6	0.6	136	2	E97963	hypothetical prote
744	6	0.6	137	2	G86789	hypothetical prote
745	6	0.6	137	2	AC1817	hypothetical prote
746	6	0.6	137	2	C87332	hypothetical prote
747	6	0.6	137	2	AD1827	hypothetical prote
748	6	0.6	138	2	AB1306	transcription regu
749	6	0.6	139	1	H64005	conserved hypothet
750	6	0.6	140	1	HBFGRE	hemoglobin beta ch
751	6	0.6	140	2	PH0134	Ig lambda chain pr
752	6	0.6	140	2	D97722	50S ribosomal prot
753	6	0.6	140	2	F70515	probable lipoprote
754	6	0.6	140	2	E71871	DNA transfer prote
755	6	0.6	140	2	C97204	probable acetyltra
756	6	0.6	140	2	AB1678	transcription regu
757	6	0.6	141	1	HAPN	hemoglobin alpha-A
758	6	0.6	141	1	HAGSM	hemoglobin alpha-A
759	6	0.6	141	1	HADKAM	hemoglobin alpha-A
760	6	0.6	141	1	HADKAW	hemoglobin alpha-A
761	6	0.6	141	1	HAGS	hemoglobin alpha-A
762	6	0.6	141	1	HAGSC	hemoglobin alpha-A
763	6	0.6	141	1	HAWS	hemoglobin alpha-A
764	6	0.6	141	1	HAGSAA	hemoglobin alpha-A
765	6	0.6	141	1	HACOAG	hemoglobin alpha-A
766	6	0.6	141	1	HAGLAB	hemoglobin alpha-A
767	6	0.6	141	1	HADLA	hemoglobin alpha-A
768	6	0.6	141	1	HAJSA	hemoglobin alpha-A
769	6	0.6	141	1	HAHTAB	hemoglobin alpha-A
770	6	0.6	141	1	HAIG1	hemoglobin alpha-I
771	6	0.6	141	1	HADKDD	hemoglobin alpha-D
772	6	0.6	141	1	HADKMD	hemoglobin alpha-D
773	6	0.6	141	1	HAGSDA	hemoglobin alpha-D
774	6	0.6	141	1	HAGSD	hemoglobin alpha-D
775	6	0.6	141	1	HAGSDI	hemoglobin alpha-D
776	6	0.6	141	1	HAGSDC	hemoglobin alpha-D
777	6	0.6	141	2	A24625	hemoglobin alpha-A
778	6	0.6	141	2	A40463	integrin alpha-6 c
779	6	0.6	141	2	C69018	methyl viologen-re
780	6	0.6	141	2	A30315	methyl viologen-re
781	6	0.6	141	2	A44464	oxaloacetate decar
782	6	0.6	141	2	C70572	hypothetical prote
783	6	0.6	141	2	F69948	phage-related prot
784	6	0.6	142	1	HART1	hemoglobin alpha-1
785	6	0.6	142	1	HADK	hemoglobin alpha-A
786	6	0.6	142	1	HADKAY	hemoglobin alpha-A
787	6	0.6	142	2	S39908	S-receptor kinase
788	6	0.6	142	2	S13452	hemoglobin alpha-A
789	6	0.6	142	2	A44777	profilin spCoel1 -
790	6	0.6	142	2	B65049	hypothetical prote
791	6	0.6	142	2	G91072	hypothetical prote
792	6	0.6	142	2	B85917	hypothetical prote
793	6	0.6	142	2	AC1642	hypothetical prote
794	6	0.6	142	2	G87647	hypothetical prote
795	6	0.6	142	2	F72475	hypothetical prote
796	6	0.6	143	1	WZBE32	gene 32 protein -
797	6	0.6	143	2	S47832	hypothetical 15.6K
798	6	0.6	143	2	A98190	hypothetical prote
799	6	0.6	143	2	B86037	hypothetical prote
800	6	0.6	143	2	AG0974	probable secreted
801	6	0.6	143	2	D97017	ferric uptake regu
802	6	0.6	143	2	S66915	hypothetical prote
803	6	0.6	143	2	D81707	hypothetical prote
804	6	0.6	143	2	H70424	pilin - Aquifex ae
805	6	0.6	143	2	T23735	hypothetical prote
806	6	0.6	143	2	F95116	GtrA family protei
807	6	0.6	143	2	D97986	mesH proteine [imp
808	6	0.6	144	2	JN0480	phospholipase A2 (
809	6	0.6	144	2	AF2292	hypothetical prote
810	6	0.6	145	1	PSBOA	phospholipase A2 (
811	6	0.6	145	1	TPQJT1	troponin T, embryo
812	6	0.6	145	1	TPQJT2	troponin T, embryo
813	6	0.6	145	2	E70168	flagellar protein
814	6	0.6	145	2	F82189	hypothetical prote
815	6	0.6	145	2	D84407	iron-dependent rep
816	6	0.6	146	1	C34588	DNA-directed RNA p
817	6	0.6	146	2	D65134	hypothetical 16.8
818	6	0.6	146	2	C91158	hypothetical prote
819	6	0.6	146	2	A86004	hypothetical prote
820	6	0.6	146	2	B64950	copper homeostasis
821	6	0.6	146	2	H90951	copper homeostasis
822	6	0.6	146	2	D85800	copper homeostasis
823	6	0.6	146	2	AI0733	probable bacteriop
824	6	0.6	146	2	F69026	conserved hypothet
825	6	0.6	147	2	F82735	hypothetical prote
826	6	0.6	147	2	E82523	hypothetical prote
827	6	0.6	147	2	S73443	hypothetical prote
828	6	0.6	148	2	T23839	hypothetical prote
829	6	0.6	148	2	B69813	hypothetical prote
830	6	0.6	148	2	E69032	formate hydrogenly
831	6	0.6	148	2	T39949	probable copper tr
832	6	0.6	149	2	S39910	S-receptor kinase

833	6	0.6	149	2	S31391	interleukin-2 prec
834	6	0.6	149	2	B40463	integrin alpha-6 c
835	6	0.6	149	2	T37099	hypothetical prote
836	6	0.6	149	2	F70599	hypothetical prote
837	6	0.6	149	2	S73351	adhesin P1 homolog
838	6	0.6	149	2	B86475	hypothetical prote
839	6	0.6	149	2	S22209	photosystem I chai
840	6	0.6	149	2	T05934	probable jasmonate
841	6	0.6	150	2	S31411	S-receptor kinase-
842	6	0.6	150	2	F65100	hypothetical 16.3
843	6	0.6	150	2	A91128	hypothetical prote
844	6	0.6	150	2	H85972	hypothetical prote
845	6	0.6	150	2	S58172	mithramycin polyke
846	6	0.6	150	2	S78334	photosystem I chai
847	6	0.6	150	2	T47315	hypothetical prote
848	6	0.6	151	2	F89973	conserved hypothet
849	6	0.6	152	2	S00332	alpha-amylase inhi
850	6	0.6	152	2	C75544	hypothetical prote
851	6	0.6	152	2	T24064	hypothetical prote
852	6	0.6	153	1	ICHU2	interleukin-2 prec
853	6	0.6	153	1	ICGI2	interleukin-2 prec
854	6	0.6	153	2	F75514	conserved hypothet
855	6	0.6	153	2	S47817	hypothetical 18.1K
856	6	0.6	153	2	G91187	hypothetical prote
857	6	0.6	153	2	H86034	hypothetical prote
858	6	0.6	153	2	F85642	hypothetical prote
859	6	0.6	153	2	A90782	hypothetical prote
860	6	0.6	153	2	AF3648	hypothetical prote
861	6	0.6	153	2	G86784	hypothetical prote
862	6	0.6	153	2	S67294	hypothetical prote
863	6	0.6	154	2	JN0698	interleukin 2 prec
864	6	0.6	154	2	A69946	hypothetical prote
865	6	0.6	155	2	C64314	conserved hypothet
866	6	0.6	156	1	S06078	H+-transporting tw
867	6	0.6	156	2	S15266	fimA protein - Dic
868	6	0.6	156	2	AH2247	hypothetical prote
869	6	0.6	156	2	S31000	gene 55 protein -
870	6	0.6	156	2	D71132	hypothetical prote
871	6	0.6	157	1	S47575	EGD1 protein - yea
872	6	0.6	157	2	B70377	iron-sulfur cofact
873	6	0.6	157	2	T09526	stress response ge
874	6	0.6	157	2	T09659	pathogenesis-relat
875	6	0.6	157	2	A46566	pilin precursor -
876	6	0.6	157	2	S30129	macrogamete/sporoz
877	6	0.6	157	2	H69849	hypothetical prote
878	6	0.6	157	2	F75612	conserved hypothet
879	6	0.6	157	2	G64334	peptidylprolyl iso
880	6	0.6	157	2	AG1312	thioredoxin homolo
881	6	0.6	158	2	B54762	phospholipase A2 (
882	6	0.6	158	2	T06527	pathogenesis-relat
883	6	0.6	158	2	A75556	hypothetical prote
884	6	0.6	158	2	T09470	hypothetical prote
885	6	0.6	158	2	B85073	hypothetical prote
886	6	0.6	159	2	F82548	hypothetical prote
887	6	0.6	159	2	E83238	hypothetical prote
888	6	0.6	159	2	H86255	protein F12F1.6 [i
889	6	0.6	159	2	E90327	conserved hypothet
890	6	0.6	159	2	T04297	hypothetical prote
891	6	0.6	159	2	H95375	hypothetical prote
892	6	0.6	159	2	S49966	probable membrane
893	6	0.6	160	2	F87463	glutathione peroxi
894	6	0.6	160	2	C89832	hypothetical prote
895	6	0.6	160	2	A45607	16K surface antige
896	6	0.6	160	2	T50732	2-vinyl bacterioch
897	6	0.6	161	1	AFKTA	allophycocyanin al
898	6	0.6	161	2	S30938	allophycocyanin be
899	6	0.6	161	2	S03938	neuroendocrine pro
900	6	0.6	161	2	T21405	hypothetical prote
901	6	0.6	161	2	B72551	probable cytochrom
902	6	0.6	161	2	AG1132	hypothetical prote
903	6	0.6	162	2	T50253	Vacuolar ATP synth
904	6	0.6	162	2	C69228	ATP synthase, subu
905	6	0.6	162	2	C45731	motB protein - pha
906	6	0.6	162	2	C90952	hypothetical prote
907	6	0.6	162	2	G85800	hypothetical prote
908	6	0.6	162	2	S75277	hypothetical prote
909	6	0.6	162	2	G90958	probable tail prot
910	6	0.6	162	2	AB1950	hypothetical prote
911	6	0.6	163	1	RDLBD	dihydrofolate redu
912	6	0.6	163	2	T12561	glycine cleavage s
913	6	0.6	163	2	C71547	hypothetical prote
914	6	0.6	163	2	S58818	hypothetical prote
915	6	0.6	164	2	A85807	probable tail fibe
916	6	0.6	164	2	A89780	conserved hypothet
917	6	0.6	165	1	XUBSM1	methylated-DNA-[pr
918	6	0.6	165	2	C44938	cysteine proteinas
919	6	0.6	165	2	AF2696	hypothetical prote
920	6	0.6	165	2	F97478	ccmE (AF176798) [i
921	6	0.6	165	2	C84460	hypothetical prote
922	6	0.6	166	2	AE1212	lipoprotein signal
923	6	0.6	166	2	D90020	30S ribosomal prot
924	6	0.6	166	2	A83111	50S ribosomal prot
925	6	0.6	166	2	A28127	myosin light chain
926	6	0.6	166	2	C71874	hypothetical prote
927	6	0.6	167	1	RDSODF	dihydrofolate redu
928	6	0.6	167	2	B32646	peroxisomal membra
929	6	0.6	167	2	A32646	peroxisomal membra
930	6	0.6	167	2	S65670	phosphoprotein pho
931	6	0.6	167	2	G95335	hypothetical prote
932	6	0.6	167	2	AI1129	hypothetical prote
933	6	0.6	168	2	C96949	probable transcrip
934	6	0.6	168	2	AE0192	probable C4-dicarb
935	6	0.6	169	2	T02943	ubiquitin-conjugat
936	6	0.6	169	2	T01329	ubiquitin-conjugat
937	6	0.6	169	2	C70207	outer membrane pro
938	6	0.6	169	2	B87602	hypothetical prote
939	6	0.6	170	2	T03768	disease resistance
940	6	0.6	170	2	S59869	TRAP-like protein
941	6	0.6	170	2	E86341	F9H16.13 protein -
942	6	0.6	170	2	C86652	hypothetical prote
943	6	0.6	170	2	T49248	zinc finger-like p
944	6	0.6	171	2	T01466	iron-sulfur cofact
945	6	0.6	171	2	T27371	peptidylprolyl iso
946	6	0.6	171	2	B86719	hypothetical prote
947	6	0.6	171	2	H70527	probable PE protei
948	6	0.6	171	2	H64960	outer membrane por
949	6	0.6	171	2	AE2604	conserved hypothet
950	6	0.6	171	2	E97386	hypothetical prote
951	6	0.6	172	2	T27882	peptidylprolyl iso
952	6	0.6	172	2	T06073	peptidylprolyl iso
953	6	0.6	172	2	F84383	hypothetical prote
954	6	0.6	172	2	AF2528	hypothetical prote
955	6	0.6	172	2	F87649	ExbD/TolR family p
956	6	0.6	173	2	T27373	peptidylprolyl iso
957	6	0.6	173	2	A24859	legumin A - tick b
958	6	0.6	173	2	A72862	AcOrf-96 protein -
959	6	0.6	173	2	H86631	hypothetical prote
960	6	0.6	173	2	E42526	B6R 20.2K protein
961	6	0.6	173	2	G89852	hypothetical prote
962	6	0.6	174	2	S58492	auxin-induced prot
963	6	0.6	174	2	S12244	auxin-induced prot
964	6	0.6	174	2	G86289	auxin-induced prot
965	6	0.6	174	2	S16528	hypothetical prote
966	6	0.6	174	2	D64150	hypothetical prote
967	6	0.6	174	2	F70635	hypothetical prote
968	6	0.6	175	2	T05380	hypothetical prote
969	6	0.6	175	2	T11504	NADH2 dehydrogenas
970	6	0.6	175	2	C75300	hypothetical prote
971	6	0.6	175	2	B70762	probable fadD11' p
972	6	0.6	176	2	I84638	pituitary adenylat
973	6	0.6	176	2	S30199	ribosomal protein
974	6	0.6	177	2	S23505	chymase (EC 3.4.21
975	6	0.6	177	2	S57399	small heat shock p
976	6	0.6	177	2	C97680	hypothetical prote
977	6	0.6	178	2	S31768	cystathionine beta
978	6	0.6	178	2	AB0870	outer membrane pro

979	6	0.6	178	2	A70227	pyrazinamidase/nic
980	6	0.6	179	2	G97729	ubiquinone biosynt
981	6	0.6	179	2	S61904	hypothetical prote
982	6	0.6	179	2	S74360	hypothetical prote
983	6	0.6	179	2	B70757	hypothetical prote
984	6	0.6	179	2	H86273	F7A19.14 protein -
985	6	0.6	180	2	H95095	hypothetical prote
986	6	0.6	181	2	D69475	conserved hypothet
987	6	0.6	181	2	I69096	cdtC protein - Esc
988	6	0.6	181	2	T18074	hypothetical prote
989	6	0.6	181	2	AI2402	hypothetical prote
990	6	0.6	181	2	E86488	protein T32E20.35
991	6	0.6	181	2	S74786	hypothetical prote
992	6	0.6	181	2	E96655	hypothetical prote
993	6	0.6	182	1	RXPG	relaxin precursor
994	6	0.6	182	2	F81366	adenine phosphorib
995	6	0.6	182	2	C83600	hypothetical prote
996	6	0.6	182	2	A84749	hypothetical prote
997	6	0.6	182	2	C71214	hypothetical prote
998	6	0.6	183	2	A28335	leukocyte common a
999	6	0.6	183	2	A65092	hypothetical fimbr
1000	6	0.6	183	2	H71894	probable flagellar

ALIGNMENTS

```
RESULT 1
I41078
hemolysin - Escherichia coli
C;Species: Escherichia coli
C;Date: 31-May-1996 #sequence_revision 31-May-1996 #text_change 21-Jul-2000
C;Accession: I41078
R;Schmidt, H.; Beutin, L.; Karch, H.
Infect. Immun. 63, 1055-1061, 1995
A;Title: Molecular analysis of the plasmid-encoded hemolysin of Escherichia coli O157:H7
A;Reference number: I41077; MUID:95172699; PMID:7868227
A;Accession: I41078
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-998 <RES>
A;Cross-references: EMBL:X79839; NID:g860924; PIDN:CAA56234.1; PID:g4388764
C;Superfamily: hemolysin A; hemolysin A homology
C;Keywords: lipoprotein
F;233-776/Domain: hemolysin A homology <HLYA>
F;550,675/Binding site: palmitate (Lys) (covalent) #status predicted

  Query Match             1.5%;  Score 14;  DB 2;  Length 998;
  Best Local Similarity   100.0%;  Pred. No. 6.2e-05;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        379 ISGILEASKQAMFE 392
              ||||||||||||||
Db        392 ISGILEASKQAMFE 405


RESULT 2
T00227
hemolysin A toxin protein - Escherichia coli plasmid pO157
C;Species: Escherichia coli
C;Date: 01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change 21-Jul-2000
C;Accession: T00227; T42148
R;Makino, K.; Ishii, K.; Yasunaga, T.; Hattori, M.; Yokoyama, K.; Yatsudo, H.C.; Kubota,
 S.; Shinagawa, H.
DNA Res. 5, 1-9, 1998
A;Title: Complete nucleotide sequences of 93-kb and 3.3-kb plasmids of an enterohemorrha
A;Reference number: Z14127; MUID:98290540; PMID:9628576
A;Accession: T00227
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-998 <MAK>
A;Cross-references: EMBL:AB011549; NID:g4589740; PIDN:BAA31774.1; PID:g3337015
A;Experimental source: strain EHEC O157:H7, substrain RIMD 0509952
R;Burland, V.; Shao, Y.; Perna, N.T.; Plunkett, G.; Sofia, H.J.; Blattner, F.R.
Nucleic Acids Res. 26, 4196-4204, 1998
A;Title: The complete DNA sequence and analysis of the large virulence plasmid of Escher
A;Reference number: Z22068; MUID:98391744; PMID:9722640
A;Accession: T42148
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-998 <BUR>
A;Cross-references: EMBL:AF074613; PIDN:AAC70116.1
A;Experimental source: strain EDL933; serotype O157:H7
C;Genetics:
A;Gene: hlyA
A;Genome: plasmid pO157
C;Superfamily: hemolysin A; hemolysin A homology
C;Keywords: cytolysis; hemolysis; lipoprotein; toxin
F;233-776/Domain: hemolysin A homology <HLYA>
F;550,675/Binding site: palmitate (Lys) (covalent) #status predicted

  Query Match             1.5%;  Score 14;  DB 2;  Length 998;
  Best Local Similarity   100.0%;  Pred. No. 6.2e-05;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        379 ISGILEASKQAMFE 392
              ||||||||||||||
Db        392 ISGILEASKQAMFE 405


RESULT 3
LEECA
hemolysin A - Escherichia coli
C;Species: Escherichia coli
C;Date: 30-Sep-1988 #sequence_revision 30-Sep-1988 #text_change 18-Jun-1999
C;Accession: A24433; I41280
R;Felmlee, T.; Pellett, S.; Welch, R.A.
J. Bacteriol. 163, 94-105, 1985
A;Title: Nucleotide sequence of an Escherichia coli chromosomal hemolysin.
A;Reference number: A24433; MUID:85234404; PMID:3891743
A;Accession: A24433
A;Molecule type: DNA
A;Residues: 1-1023 <FEL>
A;Cross-references: GB:M10133; GB:M12863; NID:g146377; PIDN:AAA23975.1; PID:g146379
A;Experimental source: strain J96, O4 serotype
R;Stanley, P.; Packman, L.C.; Koronakis, V.; Hughes, C.
Science 266, 1992-1996, 1994
A;Title: Fatty acylation of two internal lysine residues required for the toxic activity
A;Reference number: A55387; MUID:95099325; PMID:7801126
A;Contents: annotation; lysine palmitoylation
A;Note: lysine modification is performed by the hlyC gene product
R;Haertlein, M.; Schiessl, S.; Wagner, W.; Rdest, U.; Kreft, J.; Goebel, W.
J. Cell Biol. 22, 87-97, 1983
A;Title: Transport of hemolysin by Escherichia coli.
A;Reference number: I41280
A;Accession: I41280
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1,'T',3,'V',5,'T',7-44 <RES>
A;Cross-references: GB:M29173; NID:g146337; PIDN:AAA23957.1; PID:g146338
C;Genetics:
A;Gene: hlyA
C;Function:
A;Description: attacks blood cell membranes and causes cell lysis
C;Superfamily: hemolysin A; hemolysin A homology
C;Keywords: calcium binding; cytolysis; exotoxin; hemolysis; lipoprotein; tandem repeat;
F;246-791/Domain: hemolysin A homology <HLYA>
F;723-851/Region: 9-residue repeats (G-G-X-G-[DN]-D-X-[LVIYF]-X)
F;563,689/Binding site: palmitate (Lys) (covalent) #status experimental

  Query Match             1.5%;  Score 14;  DB 1;  Length 1023;
  Best Local Similarity   100.0%;  Pred. No. 6.4e-05;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        379 ISGILEASKQAMFE 392
              ||||||||||||||
```

```
Db        405 ISGILEASKQAMFE 418


RESULT 4
S10056
hemolysin A - Escherichia coli plasmid pHly152
C;Species: Escherichia coli
C;Date: 19-Mar-1997 #sequence_revision 25-Apr-1997 #text_change 04-Mar-2000
C;Accession: S10056
R;Hess, J.; Wels, W.; Vogel, M.; Goebel, W.
FEMS Microbiol. Lett. 34, 1-11, 1986
A;Title: Nucleotide sequence of a plasmid-encoded hemolysin determinant and its comparis
A;Reference number: S07209
A;Accession: S10056
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1024 <HES>
A;Cross-references: EMBL:M14107
C;Genetics:
A;Genome: plasmid pHly152
C;Superfamily: hemolysin A; hemolysin A homology
C;Keywords: lipoprotein
F;247-792/Domain: hemolysin A homology <HLYA>
F;564,690/Binding site: palmitate (Lys) (covalent) #status predicted

  Query Match             1.5%;  Score 14;  DB 2;  Length 1024;
  Best Local Similarity   100.0%;  Pred. No. 6.4e-05;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        379 ISGILEASKQAMFE 392
              ||||||||||||||
Db        406 ISGILEASKQAMFE 419


RESULT 5
S51784
toxin III - Actinobacillus pleuropneumoniae (serotype 2)
N;Alternate names: RTX-toxin IIIA (ApxIIIA)
C;Species: Actinobacillus pleuropneumoniae
C;Date: 14-Jul-1995 #sequence_revision 15-Nov-1996 #text_change 18-Jun-1999
C;Accession: S51784
R;Chang, Y.F.; Shi, J.; Ma, D.P.; Shin, S.J.; Lein, D.H.
DNA Cell Biol. 12, 351-362, 1993
A;Title: Molecular analysis of the Actinobacillus pleuropneumoniae RTX toxin-III gene cl
A;Reference number: S51783; MUID:93263992; PMID:8494611
A;Accession: S51784
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1049 <CHA>
A;Cross-references: EMBL:L12145; NID:g349605; PIDN:AAA21924.1; PID:g470685
C;Comment: This orgnism causes porcine pleuropneumonia.
C;Genetics:
A;Gene: apxIIIA
C;Function:
A;Description: lyses lung macrophages
C;Superfamily: hemolysin A; hemolysin A homology
C;Keywords: calcium binding; cytolysis; exotoxin; lipoprotein; tandem repeat; thiolester
F;254-803/Domain: hemolysin A homology <HLYA>
F;735-861/Region: 9-residue repeats (G-G-X-G-[DN]-D-X-[LVIYF]-X)
F;571,702/Binding site: palmitate (Lys) (covalent) #status predicted

  Query Match             1.5%;  Score 14;  DB 1;  Length 1049;
  Best Local Similarity   100.0%;  Pred. No. 6.5e-05;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        241 KVAAGFELSNQVIG 254
              ||||||||||||||
Db        275 KVAAGFELSNQVIG 288


RESULT 6
B49219
toxin III - Actinobacillus pleuropneumoniae (serotype 8)
N;Alternate names: RTX-toxin III (ApxIIIA)
C;Species: Actinobacillus pleuropneumoniae
C;Date: 19-Dec-1993 #sequence_revision 01-Nov-1996 #text_change 18-Jun-1999
C;Accession: B49219; S48043; S29958
R;Jansen, R.; Briaire, J.; Kamp, E.M.; Gielkens, A.L.; Smits, M.A.
Infect. Immun. 61, 947-954, 1993
A;Title: Cloning and characterization of the Actinobacillus pleuropneumoniae-RTX-toxin I]
A;Reference number: A49219; MUID:93162836; PMID:8432615
A;Accession: B49219
A;Molecule type: DNA
A;Residues: 1-1052 <JAN1>
A;Cross-references: EMBL:X68815; NID:g38956; PIDN:CAA48711.1; PID:g38958
A;Experimental source: strain 405, serotype 8
A;Note: sequence extracted from NCBI backbone (NCBIN:125168, NCBIP:125170)
R;Jansen, R.; Briaire, J.; van Geel, A.B.M.; Kamp, E.M.; Gielkens, A.L.J.; Smits, M.A.
Infect. Immun. 62, 4411-4418, 1994
A;Title: Genetic map of the Actinobacillus pleuropneumoniae RTX-toxin (Apx) operons: cha
A;Reference number: S48042; MUID:95012630; PMID:7927703
A;Accession: S48043
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-1052 <JAN2>
A;Cross-references: EMBL:X80055; NID:g558150; PIDN:CAA56358.1; PID:g558152
A;Experimental source: strain 405, serotype 8
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, July 1994
C;Comment: This orgnism causes porcine pleuropneumonia.
C;Genetics:
A;Gene: apxIIIA
C;Function:
A;Description: lyses lung macrophages
C;Superfamily: hemolysin A; hemolysin A homology
C;Keywords: calcium binding; cytolysis; exotoxin; lipoprotein; tandem repeat; thiolester
F;254-804/Domain: hemolysin A homology <HLYA>
F;736-862/Region: 9-residue repeats (G-G-X-G-[DN]-D-X-[LVIYF]-X)
F;571,702/Binding site: palmitate (Lys) (covalent) #status predicted

  Query Match             1.5%;  Score 14;  DB 1;  Length 1052;
  Best Local Similarity   100.0%;  Pred. No. 6.5e-05;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        241 KVAAGFELSNQVIG 254
              ||||||||||||||
Db        275 KVAAGFELSNQVIG 288


RESULT 7
B33389
toxin II - Actinobacillus pleuropneumoniae
N;Alternate names: cytolysin II; RTX-toxin II (ApxII)
C;Species: Actinobacillus pleuropneumoniae
C;Date: 09-Mar-1990 #sequence_revision 01-Nov-1996 #text_change 18-Jun-1999
C;Accession: B33389; S18853; B43599
R;Chang, Y.F.; Young, R.; Struck, D.K.
DNA 8, 635-647, 1989
A;Title: Cloning and characterization of a hemolysin gene from Actinobacillus (Haemophil
A;Reference number: A33389; MUID:90126233; PMID:2693022
A;Accession: B33389
A;Molecule type: DNA
A;Residues: 1-956 <CHA>
A;Cross-references: GB:M30602; NID:g141823; PIDN:AAA87232.1; PID:g141825
A;Experimental source: serotype 5
R;Smits, M.A.; Briaire, J.; Jansen, R.; Smith, H.E.; Kamp, E.M.; Gielkens, A.L.J.
submitted to the EMBL Data Library, July 1991
A;Description: Cytolysins of Actinobacillus pleuropneumoniae serotype 9.
A;Reference number: S18852
A;Accession: S18853
A;Molecule type: DNA
A;Residues: 1-956 <SMI>
A;Cross-references: EMBL:X61111; NID:g38939; PIDN:CAA43423.1; PID:g38941
R;Smits, M.A.; Briaire, J.; Jansen, R.; Smith, H.E.; Kamp, E.M.; Gielkens, A.L.J.
Infect. Immun. 59, 4497-4504, 1991
A;Title: Cytolysins of Actinobacillus pleuropneumoniae serotype 9.
A;Reference number: A43599; MUID:92040145; PMID:1937809
```

```
A;Accession: B43599
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-27;948-956 <SM2>
A;Cross-references: GB:X61111; NID:g38939
C;Comment: This organism causes porcine pleuropneumonia.
C;Genetics:
A;Gene: apxIIA; appA; clyIIA
C;Function:
A;Description: attacks blood cell membranes and causes cell lysis
C;Superfamily: hemolysin A; hemolysin A homology
C;Keywords: calcium binding; cytolysis; exotoxin; hemolysis; lipoprotein; tandem repeat;
F;243-787/Domain: hemolysin A homology <HLYA>
F;719-801/Region: 9-residue repeats (G-G-X-G-[DN]-D-X-[LVIYF]-X)
F;557/Binding site: palmitate (Lys) (covalent) #status predicted

  Query Match              1.2%;  Score 11;  DB 1;  Length 956;
  Best Local Similarity  100.0%;  Pred. No. 0.062;
  Matches   11;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        369 ALLVAGVTGLI 379
              |||||||||||
Db        392 ALLVAGVTGLI 402


RESULT 8
A43834
toxin II - Actinobacillus suis
N;Alternate names: ashA protein; cytolysin II; RTX-toxin II
C;Species: Actinobacillus suis
C;Date: 31-Dec-1993 #sequence_revision 08-Nov-1996 #text_change 05-Dec-1998
C;Accession: A43834
R;Burrows, L.L.; Lo, R.Y.
Infect. Immun. 60, 2166-2173, 1992
A;Title: Molecular characterization of an RTX toxin determinant from Actinobacillus suis
A;Reference number: A43834; MUID:92267623; PMID:1587585
A;Accession: A43834
A;Molecule type: DNA
A;Residues: 1-956 <BUR>
A;Experimental source: isolate 3714
A;Note: sequence extracted from NCBI backbone (NCBIN:104212, NCBIP:104211)
C;Comment: This organism causes acute fatal septicemia in young pigs.
C;Function:
A;Description: attacks cell membranes and causes cell lysis
C;Superfamily: hemolysin A; hemolysin A homology
C;Keywords: calcium binding; cytolysis; exotoxin; hemolysis; lipoprotein; tandem repeat;
F;243-787/Domain: hemolysin A homology <HLYA>
F;719-801/Region: 9-residue repeats (G-G-X-G-[DN]-D-X-[LVIYF]-X)
F;557/Binding site: palmitate (Lys) (covalent) #status predicted

  Query Match              1.2%;  Score 11;  DB 1;  Length 956;
  Best Local Similarity  100.0%;  Pred. No. 0.062;
  Matches   11;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        369 ALLVAGVTGLI 379
              |||||||||||
Db        392 ALLVAGVTGLI 402


RESULT 9
I39643
RTX-toxin I - Actinobacillus pleuropneumoniae
N;Alternate names: hemolysin ApxI
C;Species: Actinobacillus pleuropneumoniae
C;Date: 19-Jul-1996 #sequence_revision 08-Nov-1996 #text_change 18-Jun-1999
C;Accession: I39643; S18769; I39645; S60732; S35781
R;Jansen, R.; Briaire, J.; Kamp, E.M.; Gielkens, A.L.; Smits, M.A.
Infect. Immun. 61, 3688-3695, 1993
A;Title: Structural analysis of the Actinobacillus pleuropneumoniae-RTX-toxin I (ApxI) c
A;Reference number: I39641; MUID:93366425; PMID:8359891
A;Accession: I39643
A;Status: nucleic acid sequence not shown; translation not shown; translated from GB/EMB
A;Molecule type: DNA
A;Residues: 1-1022 <RES>
A;Cross-references: EMBL:X73117; NID:g312897; PIDN:CAA51548.1; PID:g312899
R;Frey, J.; Meier, R.; Gygi, D.; Nicolet, J.
Infect. Immun. 59, 3026-3032, 1991
A;Title: Nucleotide sequence of the hemolysin I gene from Actinobacillus pleuropneumonia
A;Reference number: S18769; MUID:91348845; PMID:1879928
A;Accession: S18769
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-209,'AMPYLTLA',218-373,'R',375-561,'Q',563-686,'TC',688-1022 <FRE>
A;Cross-references: EMBL:X52899; NID:g38949; PIDN:CAA37081.1; PID:g38950
R;Frey, J.; Haldimann, A.; Nicolet, J.; Boffini, A.; Prentki, P.
Gene 142, 97-102, 1994
A;Title: Sequence analysis and transcription of the apxI operon (hemolysin I) from Actin
A;Reference number: I39644; MUID:94237497; PMID:8181764
A;Accession: I39645
A;Status: nucleic acid sequence not shown; translation not shown; translated from GB/EMBI
A;Molecule type: DNA
A;Residues: 1-209,'AMPYLTLA',218-373,'R',375-561,'Q',563-686,'TC',688-1022 <RE2>
A;Cross-references: EMBL:X68595; NID:g505568; PIDN:CAA48586.1; PID:g505570
R;Tascon, R.I.; Vazquez-Boland, J.A.; Gutierrez-Martin, C.B.; Rodriguez-Barbosa, I.; Rodr
Mol. Microbiol. 14, 207-216, 1994
A;Title: The RTX haemolysins ApxI and ApxII are major virulence factors of the swine patl
A;Reference number: S60731; MUID:95131743; PMID:7830567
A;Accession: S60732
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 604-685 <TAS>
C;Comment: This organism causes porcine pleuropneumonia.
C;Genetics:
A;Gene: apxIA
C;Function:
A;Description: attacks blood cell membranes and causes cell lysis
C;Superfamily: hemolysin A; hemolysin A homology
C;Keywords: calcium binding; cytolysis; exotoxin; hemolysis; lipoprotein; tandem repeat;
F;243-789/Domain: hemolysin A homology <HLYA>
F;721-847/Region: 9-residue repeats (G-G-X-G-[DN]-D-X-[LVIYF]-X)
F;560,686/Binding site: palmitate (Lys) (covalent) #status predicted

  Query Match              1.2%;  Score 11;  DB 1;  Length 1022;
  Best Local Similarity  100.0%;  Pred. No. 0.066;
  Matches   11;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        266 AQRVAAGLSTT 276
              |||||||||||
Db        289 AQRVAAGLSTT 299


RESULT 10
B30169
leukotoxin A - Pasteurella haemolytica (serotype 1)
N;Alternate names: lktA protein
C;Species: Pasteurella haemolytica
C;Date: 12-Oct-1989 #sequence_revision 15-Nov-1996 #text_change 18-Jun-1999
C;Accession: B30169; C32051; S29516
R;Highlander, S.K.; Chidambaram, M.; Engler, M.J.; Weinstock, G.M.
DNA 8, 15-28, 1989
A;Title: DNA sequence of the Pasteurella haemolytica leukotoxin gene cluster.
A;Reference number: A30169; MUID:89210283; PMID:2707120
A;Accession: B30169
A;Status: not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 1-953 <HIG>
R;Strathdee, C.A.; Lo, R.Y.C.
J. Bacteriol. 171, 916-928, 1989
A;Title: Cloning, nucleotide sequence, and characterization of genes encoding the secret
A;Reference number: A32051; MUID:89123172; PMID:2914876
A;Accession: C32051
A;Status: not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 947-953 <STR>
R;Lo, R.Y.C.; Strathdee, C.A.; Shewen, P.E.
Infect. Immun. 55, 1987-1996, 1987
```

```
A;Title: Nucleotide sequence of the leukotoxin genes of Pasteurella haemolytica A1.
A;Reference number: S29515; MUID:87306837; PMID:3040588
A;Accession: S29516
A;Molecule type: DNA
A;Residues: 1-741,'D',743-953 <LOR>
A;Cross-references: EMBL:M20730; NID:g150492; PIDN:AAA25529.1; PID:g150494
C;Comment: This organism causes bovine pneumonic pasteurellosis (shipping fever).
C;Genetics:
A;Gene: lktA
C;Function:
A;Description: lyses leukocytes
C;Superfamily: hemolysin A; hemolysin A homology
C;Keywords: calcium binding; cytolysis; exotoxin; hemolysis; lipoprotein; tandem repeat;
F;238-784/Domain: hemolysin A homology <HLYA>
F;716-807/Region: 9-residue repeats (G-G-X-G-[DN]-D-X-[LVIYF]-X)
F;554/Binding site: palmitate (Lys) (covalent) #status predicted

  Query Match             1.1%;  Score 10;  DB 1;  Length 953;
  Best Local Similarity   100.0%;  Pred. No. 0.62;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        265 LAQRVAAGLS 274
              ||||||||||
Db        283 LAQRVAAGLS 292


RESULT 11
A35254
leukotoxin A - Pasteurella haemolytica (serotype T10)
N;Alternate names: lktA protein
C;Species: Pasteurella haemolytica
C;Date: 10-Aug-1990 #sequence_revision 15-Nov-1996 #text_change 18-Jun-1999
C;Accession: S37145; A35254; S34237; S34235
R;Lainson, A.F.; Aitchison, K.; Donachie, W.
submitted to the EMBL Data Library, September 1993
A;Description: DNA sequence of the leukotoxin A gene from P. haemolytica T10 serotype.
A;Reference number: S37145
A;Accession: S37145
A;Molecule type: DNA
A;Residues: 1-955 <LAI>
A;Cross-references: EMBL:Z26247; NID:g400424; PIDN:CAA81206.1; PID:g400425
R;Highlander, S.K.; Engler, M.J.; Weinstock, G.M.
J. Bacteriol. 172, 2343-2350, 1990
A;Title: Secretion and expression of the Pasteurella haemolytica leukotoxin.
A;Reference number: A35254; MUID:90236888; PMID:2185213
A;Accession: A35254
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 950-955 <HIG>
A;Cross-references: GB:M24197; GB:M34943; GB:M34944
R;Lainson, A.F.; Aitchison, K.D.; Donachie, W.
submitted to the EMBL Data Library, June 1993
A;Description: DNA sequence of the carboxy terminal end of leukotoxin A from the T3 sero
A;Reference number: S34235
A;Accession: S34237
A;Molecule type: DNA
A;Residues: 745-955 <LA2>
A;Cross-references: EMBL:Z22884; NID:g311828; PIDN:CAA80498.1; PID:g311829
A;Experimental source: serotype T3
A;Accession: S34235
A;Molecule type: DNA
A;Residues: 723-955 <LA3>
A;Cross-references: EMBL:Z22887; NID:g311824; PIDN:CAA80501.1; PID:g311825
A;Experimental source: serotype T10
C;Function:
A;Description: attacks cell membranes and causes cell lysis
C;Superfamily: hemolysin A; hemolysin A homology
C;Keywords: calcium binding; cytolysis; exotoxin; hemolysis; lipoprotein; tandem repeat;
F;240-786/Domain: hemolysin A homology <HLYA>
F;718-809/Region: 9-residue repeats (G-G-X-G-[DN]-D-X-[LVIYF]-X)
F;718-726/Region: repeat
F;727-735/Region: repeat
F;736-744/Region: repeat
F;745-753/Region: repeat
F;754-762/Region: repeat
F;763-771/Region: repeat
F;772-780/Region: repeat
F;781-789/Region: repeat
F;792-800/Region: repeat
F;801-809/Region: repeat
F;556/Binding site: palmitate (Lys) (covalent) #status predicted

  Query Match             1.1%;  Score 10;  DB 1;  Length 955;
  Best Local Similarity   100.0%;  Pred. No. 0.62;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        265 LAQRVAAGLS 274
              ||||||||||
Db        285 LAQRVAAGLS 294


RESULT 12
S77624
mannuronan C-5-epimerase (EC 5.1.3.-) - Azotobacter vinelandii
C;Species: Azotobacter vinelandii
C;Date: 24-Oct-1998 #sequence_revision 24-Oct-1998 #text_change 21-Jul-2000
C;Accession: S77624
R;Ertesvaag, H.; Hoeidal, H.K.; Hals, I.K.; Rian, A.; Doseth, B.; Valla, S.
Mol. Microbiol. 16, 719-731, 1995
A;Title: A family of modular type mannuronan C-5-epimerase genes controls alginate struct
A;Reference number: I39738; MUID:96065700; PMID:7476166
A;Accession: S77624
A;Status: nucleic acid sequence not shown
A;Molecule type: DNA
A;Residues: 1-1403 <ERT>
A;Cross-references: EMBL:L39096; NID:g790690; PIDN:AAA87311.1; PID:g790692
A;Experimental source: strain E
C;Genetics:
A;Gene: algE1
C;Function:
A;Description: catalyzes the Ca(2+)-dependent epimerization of D-mannuronic acid residues
A;Pathway: alginate biosynthesis
C;Keywords: calcium binding; isomerase

  Query Match             1.1%;  Score 10;  DB 2;  Length 1403;
  Best Local Similarity   100.0%;  Pred. No. 0.88;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        714 GDDLLDGGAG 723
              ||||||||||
Db       1257 GDDLLDGGAG 1266


RESULT 13
AH2515
hypothetical protein alr7304 [imported] - Nostoc sp. (strain PCC 7120) plasmid pCC7120alp
C;Species: Nostoc sp. PCC 7120
A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002
C;Accession: AH2515
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi,
 Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S
DNA Res. 8, 205-213, 2001
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anak
A;Reference number: AB1807; MUID:21595285; PMID:11759840
A;Accession: AH2515
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-4936 <KUR>
A;Cross-references: GB:BA000020; PIDN:BAB78388.1; PID:g17135842; GSPDB:GN00180
A;Experimental source: strain PCC 7120
C;Genetics:
A;Gene: alr7304
A;Genome: plasmid

  Query Match             1.1%;  Score 10;  DB 2;  Length 4936;
```

```
  Best Local Similarity   100.0%;  Pred. No. 2.7;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        752 DVLNGGAGND 761
              ||||||||||
Db       4296 DVLNGGAGND 4305


RESULT 14
AG2664
ATP synthase B chain atpF [imported] - Agrobacterium tumefaciens (strain C58, Dupont)
C;Species: Agrobacterium tumefaciens
C;Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 18-Nov-2002
C;Accession: AG2664
R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, L
erage, G.; Gillet, W.; Grant, C.; Guenthner, D.; Kutyavin, T.; Levy, R.; Li, M.; McClell
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
ster, E.W.
A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A;Reference number: AB2577; MUID:21608550; PMID:11743193
A;Accession: AG2664
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-161 <KUR>
A;Cross-references: GB:AE008688; PIDN:AAL41733.1; PID:g17739083; GSPDB:GN00186
A;Experimental source: strain C58 (Dupont)
C;Genetics:
A;Gene: atpF
A;Map position: circular chromosome

  Query Match             0.9%;  Score 8;  DB 2;  Length 161;
  Best Local Similarity   100.0%;  Pred. No. 13;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        325 HLLAEYQR 332
              ||||||||
Db         56 HLLAEYQR 63


RESULT 15
F97446
hypothetical protein AGR_C_1301 [imported] - Agrobacterium tumefaciens (strain C58, Cere
C;Species: Agrobacterium tumefaciens
C;Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 18-Nov-2002
C;Accession: F97446
R;Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman,
 A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.;
Science 294, 2323-2328, 2001
A;Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum
A;Reference number: A97359; MUID:21608551; PMID:11743194
A;Accession: F97446
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-161 <KUR>
A;Cross-references: GB:AE007869; PIDN:AAK86527.1; PID:g15155685; GSPDB:GN00169
C;Genetics:
A;Gene: AGR_C_1301
A;Map position: circular chromosome

  Query Match             0.9%;  Score 8;  DB 2;  Length 161;
  Best Local Similarity   100.0%;  Pred. No. 13;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        325 HLLAEYQR 332
              ||||||||
Db         56 HLLAEYQR 63


RESULT 16
S34238
leukotoxin A - Pasteurella haemolytica (fragment)
N;Alternate names: lktA protein
C;Species: Pasteurella haemolytica
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 23-Mar-2001
C;Accession: S34238; S34236
R;Lainson, A.F.; Aitchison, K.D.; Donachie, W.
submitted to the EMBL Data Library, June 1993
A;Description: DNA sequence of the carboxy terminal end of leukotoxin A from the T4 sero
of Pasturella haemolytica.
A;Reference number: S34236
A;Accession: S34238
A;Molecule type: DNA
A;Residues: 1-208 <LAI>
A;Cross-references: EMBL:Z22885; NID:g311830; PIDN:CAA80499.1; PID:g311831
A;Experimental source: serotype T4
A;Accession: S34236
A;Molecule type: DNA
A;Residues: 9-208 <LA2>
A;Cross-references: EMBL:Z22886; NID:g311826; PIDN:CAA80500.1; PID:g311827
A;Experimental source: serotype T15
C;Genetics:
A;Gene: lktA
C;Superfamily: hemolysin A; hemolysin A homology
C;Keywords: calcium binding; cytolysis; exotoxin; hemolysis; lipoprotein; tandem repeat

  Query Match             0.9%;  Score 8;  DB 2;  Length 208;
  Best Local Similarity   100.0%;  Pred. No. 16;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        725 DRLFGGKG 732
              ||||||||
Db          3 DRLFGGKG 10


RESULT 17
AD0468
conserved hypothetical protein YPO3844 [imported] - Yersinia pestis (strain CO92)
C;Species: Yersinia pestis
C;Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 02-Nov-2001
C;Accession: AD0468
R;Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;
il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell,
Nature 413, 523-527, 2001
A;Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A;Reference number: AB0001; MUID:21470413; PMID:11586360
A;Accession: AD0468
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-234 <KUR>
A;Cross-references: GB:AL590842; PIDN:CAC93312.1; PID:g15981759; GSPDB:GN00175
C;Genetics:
A;Gene: YPO3844

  Query Match             0.9%;  Score 8;  DB 2;  Length 234;
  Best Local Similarity   100.0%;  Pred. No. 18;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        134 ALAGIELD 141
              ||||||||
Db          8 ALAGIELD 15


RESULT 18
E70714
hypothetical protein Rv1514c - Mycobacterium tuberculosis (strain H37RV)
C;Species: Mycobacterium tuberculosis
C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999
C;Accession: E70714
R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S
; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A;Authors: Sqares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
```

```
A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A;Reference number: A70500; MUID:98295987; PMID:9634230
A;Accession: E70714
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-262 <COL>
A;Cross-references: GB:Z79701; GB:AL123456; NID:g3261635; PIDN:CAB02028.1; PID:e264143;
A;Experimental source: strain H37Rv
C;Genetics:
A;Gene: Rv1514c

  Query Match             0.9%;  Score 8;  DB 2;  Length 262;
  Best Local Similarity   100.0%;  Pred. No. 20;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        746 DGGSGDDV 753
              ||||||||
Db         42 DGGSGDDV 49


RESULT 19
AC2987
conserved hypothetical protein Atu3500 [imported] - Agrobacterium tumefaciens (strain C5
C;Species: Agrobacterium tumefaciens
C;Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 18-Nov-2002
C;Accession: AC2987
R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, L
erage, G.; Gillet, W.; Grant, C.; Guenthner, D.; Kutyavin, T.; Levy, R.; Li, M.; McClell
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
ster, E.W.
A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A;Reference number: AB2577; MUID:21608550; PMID:11743193
A;Accession: AC2987
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-270 <KUR>
A;Cross-references: GB:AE008689; PIDN:AAL44313.1; PID:g17741903; GSPDB:GN00187
A;Experimental source: strain C58 (Dupont)
C;Genetics:
A;Gene: Atu3500
A;Map position: linear chromosome

  Query Match             0.9%;  Score 8;  DB 2;  Length 270;
  Best Local Similarity   100.0%;  Pred. No. 20;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        143 LIKKGDAA 150
              ||||||||
Db        142 LIKKGDAA 149


RESULT 20
E98296
hypothetical protein AGR_L_2657 [imported] - Agrobacterium tumefaciens (strain C58, Cere
C;Species: Agrobacterium tumefaciens
C;Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 18-Nov-2002
C;Accession: E98296
R;Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman,
 A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.;
Science 294, 2323-2328, 2001
A;Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum
A;Reference number: A97359; MUID:21608551; PMID:11743194
A;Accession: E98296
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-270 <KUR>
A;Cross-references: GB:AE007870; PIDN:AAK89895.1; PID:g15159842; GSPDB:GN00170
C;Genetics:
A;Gene: AGR_L_2657
A;Map position: linear chromosome

  Query Match             0.9%;  Score 8;  DB 2;  Length 270;
  Best Local Similarity   100.0%;  Pred. No. 20;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        143 LIKKGDAA 150
              ||||||||
Db        142 LIKKGDAA 149


RESULT 21
JE0183
chitinase (EC 3.2.1.14) 1 - cone shell (Conus tulipa)
N;Alternate names: TBC-1
C;Species: Conus tulipa (tulip cone)
C;Date: 21-Aug-1998 #sequence_revision 21-Aug-1998 #text_change 07-May-1999
C;Accession: JE0183
R;Yamagami, T.; Ishiguro, M.
Biosci. Biotechnol. Biochem. 62, 1253-1257, 1998
A;Title: Complete amino acid sequences of chitinase-1 and -2 from bulbs of genus Tulipa.
A;Reference number: JE0183; MUID:98357241; PMID:9692212
A;Accession: JE0183
A;Molecule type: protein
A;Residues: 1-275 <YAM>
C;Comment: This enzyme catalyzes the hydrolysis of beta-1,4-linked homopolymers or oligo
C;Superfamily: alcohol sulfotransferase
C;Keywords: glycosidase; hydrolase

  Query Match             0.9%;  Score 8;  DB 2;  Length 275;
  Best Local Similarity   100.0%;  Pred. No. 21;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        692 IGSQFNDV 699
              ||||||||
Db          7 IGSQFNDV 14


RESULT 22
JC7335
chitinase (EC 3.2.1.14) 1 - cone shell (Conus tulipa)
C;Species: Conus tulipa (tulip cone)
C;Date: 08-Sep-2000 #sequence_revision 08-Sep-2000 #text_change 15-Sep-2000
C;Accession: JC7335
R;Yamagami, T.; Tsutsumi, K.; Ishiguro, M.
Biosci. Biotechnol. Biochem. 64, 1394-1401, 2000
A;Title: Cloning, sequencing, and expression of the tulip bulb chitinase-1 cDNA.
A;Reference number: JC7335
A;Accession: JC7335
A;Molecule type: mRNA
A;Residues: 1-314 <YAM>
A;Cross-references: DDBJ:AB035668
C;Comment: This enzyme, a member of class III plant chitinases, which catalyzes the hydr
or protection against fungal pathogens.
C;Genetics:
A;Gene: tbc-1
C;Superfamily: alcohol sulfotransferase
C;Keywords: glycosidase; hydrolase

  Query Match             0.9%;  Score 8;  DB 2;  Length 314;
  Best Local Similarity   100.0%;  Pred. No. 23;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        692 IGSQFNDV 699
              ||||||||
Db         33 IGSQFNDV 40


RESULT 23
AD3003
hypothetical protein accA [imported] - Agrobacterium tumefaciens (strain C58, Dupont)
C;Species: Agrobacterium tumefaciens
C;Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 18-Nov-2002
C;Accession: AD3003
R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, L
```

```
erage, G.; Gillet, W.; Grant, C.; Guenthner, D.; Kutyavin, T.; Levy, R.; Li, M.; McClell
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
ster, E.W.
A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A;Reference number: AB2577; MUID:21608550; PMID:11743193
A;Accession: AD3003
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-317 <KUR>
A;Cross-references: GB:AE008689; PIDN:AAL44442.1; PID:g17742043; GSPDB:GN00187
A;Experimental source: strain C58 (Dupont)
C;Genetics:
A;Gene: accA
A;Map position: linear chromosome
C;Superfamily: acetyl-CoA carboxylase, carboxyltransferase alpha chain

  Query Match             0.9%;  Score 8;  DB 2;  Length 317;
  Best Local Similarity   100.0%;  Pred. No. 23;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        869 ELKKLADE 876
              ||||||||
Db         20 ELKKLADE 27


RESULT 24
E98280
mpsA protein [imported] - Agrobacterium tumefaciens (strain C58, Cereon)
C;Species: Agrobacterium tumefaciens
C;Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 18-Nov-2002
C;Accession: E98280
R;Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman,
 A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.;
Science 294, 2323-2328, 2001
A;Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum
A;Reference number: A97359; MUID:21608551; PMID:11743194
A;Accession: E98280
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-317 <KUR>
A;Cross-references: GB:AE007870; PIDN:AAK89767.1; PID:g15159691; GSPDB:GN00170
C;Genetics:
A;Gene: AGR_L_2394
A;Map position: linear chromosome
C;Superfamily: acetyl-CoA carboxylase, carboxyltransferase alpha chain

  Query Match             0.9%;  Score 8;  DB 2;  Length 317;
  Best Local Similarity   100.0%;  Pred. No. 23;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        869 ELKKLADE 876
              ||||||||
Db         20 ELKKLADE 27


RESULT 25
AB1054
probable membrane protein ytfF [imported] - Salmonella enterica subsp. enterica serovar
C;Species: Salmonella enterica subsp. enterica serovar Typhi
A;Note: this species has also been called Salmonella typhi
C;Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
C;Accession: AB1054
R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
, S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A;Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
A;Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov
A;Reference number: AB0502; MUID:21534947; PMID:11677608
A;Accession: AB1054
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-321 <PAR>
A;Cross-references: GB:AL513382; PIDN:CAD06879.1; PID:g16505527; GSPDB:GN00176
C;Genetics:
A;Gene: ytfF

  Query Match             0.9%;  Score 8;  DB 2;  Length 321;
  Best Local Similarity   100.0%;  Pred. No. 24;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        368 IALLVAGV 375
              ||||||||
Db        295 IALLVAGV 302


RESULT 26
C87464
lipoic acid synthase [imported] - Caulobacter crescentus
C;Species: Caulobacter crescentus
C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 10-May-2001
C;Accession: C87464
R;Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
 B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A;Title: Complete Genome Sequence of Caulobacter crescentus.
A;Reference number: A87249; MUID:21173698; PMID:11259647
A;Accession: C87464
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-325 <STO>
A;Cross-references: GB:AE005673; NID:g13423153; PIDN:AAK23711.1; GSPDB:GN00148
C;Genetics:
A;Gene: CC1735
C;Superfamily: lipoic acid synthase

  Query Match             0.9%;  Score 8;  DB 2;  Length 325;
  Best Local Similarity   100.0%;  Pred. No. 24;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        715 DDLLDGGA 722
              ||||||||
Db        135 DDLLDGGA 142


RESULT 27
E95914
probable secreted calcium-binding protein [imported] - Sinorhizobium meliloti (strain 10
C;Species: Sinorhizobium meliloti
C;Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 30-Sep-2001
C;Accession: E95914
R;Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Hernan
Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001
A;Title: The complete sequence of the 1,683-kb pSymB megaplasmid from the N2-fixing endo
A;Reference number: A95842; MUID:21396508; PMID:11481431
A;Accession: E95914
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-353 <KUR>
A;Cross-references: GB:AL591985; PIDN:CAC48981.1; PID:g15140466; GSPDB:GN00167
A;Experimental source: strain 1021, megaplasmid pSymB
R;Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,
pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;
 L.; Hyman, R.W.; Jones, T.
Science 293, 668-672, 2001
A;Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,
hebault, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.
A;Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
A;Reference number: A96039; MUID:21368234; PMID:11474104
A;Contents: annotation
C;Genetics:
A;Gene: SMb20838
A;Genome: plasmid
```

```
  Query Match             0.9%;  Score 8;  DB 2;  Length 353;
  Best Local Similarity   100.0%;  Pred. No. 26;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        745 LDGGSGDD 752
              ||||||||
Db        221 LDGGSGDD 228


RESULT 28
AF1434
AA3-600 quinol oxidase chain II [imported] - Listeria innocua (strain Clip11262)
C;Species: Listeria innocua
C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 03-Jun-2002
C;Accession: AF1434
R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A;Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma
ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,
A;Title: Comparative genomics of Listeria species.
A;Reference number: AB1077; MUID:21537279; PMID:11679669
A;Accession: AF1434
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-368 <GLA>
A;Cross-references: GB:AL592022; PIDN:CAC95246.1; PID:g16412434; GSPDB:GN00178
A;Experimental source: strain Clip11262
C;Genetics:
A;Gene: qoxA
C;Superfamily: bo-type ubiquinol oxidase chain II precursor; cytochrome-c oxidase chain
C;Keywords: copper; electron transfer; membrane-associated complex; respiratory chain

  Query Match             0.9%;  Score 8;  DB 2;  Length 368;
  Best Local Similarity   100.0%;  Pred. No. 27;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        374 GVTGLISG 381
              ||||||||
Db         15 GVTGLISG 22


RESULT 29
AF1806
AA3-600 quinol oxidase chain II [imported] - Listeria monocytogenes (strain EGD-e)
C;Species: Listeria monocytogenes
C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 03-Jun-2002
C;Accession: AF1806
R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A;Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma
ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,
A;Title: Comparative genomics of Listeria species.
A;Reference number: AB1077; MUID:21537279; PMID:11679669
A;Accession: AF1806
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-368 <GLA>
A;Cross-references: GB:NC_003210; PIDN:CAC98228.1; PID:g16409372; GSPDB:GN00177
A;Experimental source: strain EGD-e
C;Genetics:
A;Gene: qoxA
C;Superfamily: bo-type ubiquinol oxidase chain II precursor; cytochrome-c oxidase chain
C;Keywords: copper; electron transfer; membrane-associated complex; respiratory chain

  Query Match             0.9%;  Score 8;  DB 2;  Length 368;
  Best Local Similarity   100.0%;  Pred. No. 27;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        374 GVTGLISG 381
              ||||||||
Db         15 GVTGLISG 22


RESULT 30
AB3160
santhopine deaminating protein [imported] - Agrobacterium tumefaciens (strain C58, Dupont
C;Species: Agrobacterium tumefaciens
C;Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 18-Nov-2002
C;Accession: AB3160
R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, L
erage, G.; Gillet, W.; Grant, C.; Guenthner, D.; Kutyavin, T.; Levy, R.; Li, M.; McClella
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm, F
ster, E.W.
A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A;Reference number: AB2577; MUID:21608550; PMID:11743193
A;Accession: AB3160
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-380 <KUR>
A;Cross-references: GB:AE008687; PIDN:AAL45696.1; PID:g17743424; GSPDB:GN00188
A;Experimental source: strain C58 (Dupont)
C;Genetics:
A;Gene: agaE
A;Genome: plasmid

  Query Match             0.9%;  Score 8;  DB 2;  Length 380;
  Best Local Similarity   100.0%;  Pred. No. 28;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        365 GTPIALLV 372
              ||||||||
Db        230 GTPIALLV 237


RESULT 31
T19181
hypothetical protein C10C5.5 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C;Accession: T19181
R;Matthews, P.
submitted to the EMBL Data Library, December 1995
A;Reference number: Z19085
A;Accession: T19181
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-394 <WIL>
A;Cross-references: EMBL:Z68214; PIDN:CAA92446.1; GSPDB:GN00022; CESP:C10C5.5
A;Experimental source: clone C10C5
C;Genetics:
A;Gene: CESP:C10C5.5
A;Map position: 4
A;Introns: 27/1; 48/3; 215/3; 268/1; 303/3; 330/2; 347/3

  Query Match             0.9%;  Score 8;  DB 2;  Length 394;
  Best Local Similarity   100.0%;  Pred. No. 28;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         48 ADELGIAR 55
              ||||||||
Db         37 ADELGIAR 44


RESULT 32
T19180
hypothetical protein C10C5.3 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C;Accession: T19180
```

R;Matthews, P.
submitted to the EMBL Data Library, December 1995
A;Reference number: Z19085
A;Accession: T19180
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-399 <WIL>
A;Cross-references: EMBL:Z68214; PIDN:CAA92445.1; GSPDB:GN00022; CESP:C10C5.3
A;Experimental source: clone C10C5
C;Genetics:
A;Gene: CESP:C10C5.3
A;Map position: 4
A;Introns: 27/1; 48/3; 215/3; 268/1; 303/3; 330/2; 350/3; 383/1

```
  Query Match             0.9%;  Score 8;  DB 2;  Length 399;
  Best Local Similarity   100.0%;  Pred. No. 29;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         48 ADELGIAR 55
              ||||||||
Db         37 ADELGIAR 44
```

RESULT 33
E64242
GTP-binding protein obg - Mycoplasma genitalium
C;Species: Mycoplasma genitalium
C;Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 19-Jan-2001
C;Accession: E64242
R;Fraser, C.M.; Gocayne, J.D.; White, O.; Adams, M.D.; Clayton, R.A.; Fleischmann, R.D.;
M.; Fuhrmann, J.; Nguyen, D.; Utterback, T.R.; Saudek, D.M.; Phillips, C.A.; Merrick, J.
, C.A.; Venter, J.C.
Science 270, 397-403, 1995
A;Title: The minimal gene complement of Mycoplasma genitalium.
A;Reference number: A64200; MUID:96026346; PMID:7569993
A;Accession: E64242
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-433 <TIGR>
A;Cross-references: GB:U39723; GB:L43967; NID:g1046092; PID:g1046095; TIGR:MG384
A;Experimental source: strain G-37
C;Genetics:
A;Genetic code: SGC3
C;Superfamily: GTP-binding protein obg; translation elongation factor Tu homology
C;Keywords: GTP binding; nucleotide binding; P-loop
F;160-285/Domain: translation elongation factor Tu homology <ETU>
F;166-173/Region: nucleotide-binding motif A (P-loop)
F;189-194/Region: GTP-binding #status predicted
F;212-215/Region: GTP-binding #status predicted
F;282-285/Region: GTP-binding #status predicted
F;309-313/Region: GTP-binding #status predicted

```
  Query Match             0.9%;  Score 8;  DB 1;  Length 433;
  Best Local Similarity   100.0%;  Pred. No. 31;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         90 KLEKFLQK 97
              ||||||||
Db        295 KLEKFLQK 302
```

RESULT 34
AE0826
probable cadaverine/lysine antiporter [imported] - Salmonella enterica subsp. enterica s
C;Species: Salmonella enterica subsp. enterica serovar Typhi
A;Note: this species has also been called Salmonella typhi
C;Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
C;Accession: AE0826
R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
, S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A;Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
A;Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov
A;Reference number: AB0502; MUID:21534947; PMID:11677608
A;Accession: AE0826
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-443 <PAR>
A;Cross-references: GB:AL513382; PIDN:CAD02761.1; PID:g16503771; GSPDB:GN00176
C;Genetics:
A;Gene: STY2805
C;Superfamily: L-lysine transport protein

```
  Query Match             0.9%;  Score 8;  DB 2;  Length 443;
  Best Local Similarity   100.0%;  Pred. No. 32;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        131 LGTALAGI 138
              ||||||||
Db        226 LGTALAGI 233
```

RESULT 35
C84306
hypothetical protein trp6 [imported] - Halobacterium sp. NRC-1
C;Species: Halobacterium sp. NRC-1
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 16-Feb-2001
C;Accession: C84306
R;Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S
; Leithauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablo
Jung, K.H.; Alam, M.; Freitas, T.
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
A;Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Li
A;Title: Genome sequence of Halobacterium species NRC-1.
A;Reference number: A84160; MUID:20504483; PMID:11016950
A;Accession: C84306
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-447 <STO>
A;Cross-references: GB:AE004437; NID:g10581015; PIDN:AAG19815.1; GSPDB:GN00138
C;Genetics:
A;Gene: trp6
C;Superfamily: gamma-aminobutyric acid transporter

```
  Query Match             0.9%;  Score 8;  DB 2;  Length 447;
  Best Local Similarity   100.0%;  Pred. No. 32;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        358 AAVGSAVG 365
              ||||||||
Db         16 AAVGSAVG 23
```

RESULT 36
H95369
EglC ENDO-1,3-1,4-BETA-GLYCANASE (EC 3.2.1.-) [imported] - Sinorhizobium meliloti (strai
C;Species: Sinorhizobium meliloti
C;Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 30-Sep-2001
C;Accession: H95369
R;Barnett, M.J.; Fisher, R.F.; Jones, T.; Komp, C.; Abola, A.P.; Barloy-Hubler, F.; Bows
.; Kalman, S.; Keating, D.H.; Palm, C.; Peck, M.C.; Surzycki, R.; Wells, D.H.; Yeh, K.C.
Proc. Natl. Acad. Sci. U.S.A. 98, 9883-9888, 2001
A;Title: Nucleotide sequence and predicted functions of the entire Sinorhizobium melilot
A;Reference number: A95262; MUID:21396509; PMID:11481432
A;Accession: H95369
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-465 <KUR>
A;Cross-references: GB:AE006469; PIDN:AAK65522.1; PID:g14523997; GSPDB:GN00165
A;Experimental source: strain 1021, megaplasmid pSymA
R;Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,
pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;
L.; Hyman, R.W.; Jones, T.
Science 293, 668-672, 2001
A;Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,

```
hebault, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.
A;Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
A;Reference number: A96039; MUID:21368234; PMID:11474104
A;Contents: annotation
C;Genetics:
A;Gene: eglC
A;Genome: plasmid
C;Keywords: glycosidase; hydrolase

  Query Match             0.9%;  Score 8;  DB 2;  Length 465;
  Best Local Similarity   100.0%;  Pred. No. 33;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        755 NGGAGNDV 762
              ||||||||
Db        123 NGGAGNDV 130


RESULT 37
B84372
DNA damage-inducible protein [imported] - Halobacterium sp. NRC-1
C;Species: Halobacterium sp. NRC-1
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C;Accession: B84372
R;Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S
; Leithauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablo
 Jung, K.H.; Alam, M.; Freitas, T.
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
A;Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Li
A;Title: Genome sequence of Halobacterium species NRC-1.
A;Reference number: A84160; MUID:20504483; PMID:11016950
A;Accession: B84372
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-469 <STO>
A;Cross-references: GB:AE004437; NID:g10581629; PIDN:AAG20342.1; GSPDB:GN00138
C;Genetics:
A;Gene: dinF

  Query Match             0.9%;  Score 8;  DB 2;  Length 469;
  Best Local Similarity   100.0%;  Pred. No. 33;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        357 AAAVGSAV 364
              ||||||||
Db        296 AAAVGSAV 303


RESULT 38
S60902
CDP-ribitol pyrophosphorylase - Haemophilus influenzae
C;Species: Haemophilus influenzae
C;Date: 27-Apr-1996 #sequence_revision 17-May-1996 #text_change 08-Oct-1999
C;Accession: S60902; S49238; S44071
R;van Eldere, J.; Brophy, L.; Loynds, B.; Celis, P.; Hancock, I.; Carman, S.; Kroll, J.S
Mol. Microbiol. 15, 107-118, 1995
A;Title: Region II of the Haemophilus influenzae type b capsulation locus is involved in
A;Reference number: S60902; MUID:95272382; PMID:7752885
A;Accession: S60902
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-474 <VAN>
A;Cross-references: EMBL:X78559; NID:g471233; PIDN:CAA55303.1; PID:g471234
A;Experimental source: serotype b; strain RM135
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, April 1994
R;Celis, P.; Moxon, E.; Eyssen, H.; van Eldere, J.
submitted to the EMBL Data Library, September 1994
A;Description: Genetic analysis of the region II of the Haemophilus influenzae serotype
A;Reference number: S49238
A;Accession: S49238
A;Molecule type: DNA
A;Residues: 1,'L',3-14,'I',16-39,'IF',42-70,'AG',73-101,'R',103-212,'F',214-272,'V',274-
A;Cross-references: EMBL:Z37516; NID:g547510; PIDN:CAA85750.1; PID:g547511
A;Experimental source: serotype a; strain RM107; ATCC 9006

  Query Match             0.9%;  Score 8;  DB 2;  Length 474;
  Best Local Similarity   100.0%;  Pred. No. 34;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        560 SKLDFSKV 567
              ||||||||
Db         69 SKLDFSKV 76


RESULT 39
D84306
sodium- and chloride-dependent transporter [imported] - Halobacterium sp. NRC-1
C;Species: Halobacterium sp. NRC-1
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 16-Feb-2001
C;Accession: D84306
R;Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S.
; Leithauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablor
 Jung, K.H.; Alam, M.; Freitas, T.
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
A;Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Lia
A;Title: Genome sequence of Halobacterium species NRC-1.
A;Reference number: A84160; MUID:20504483; PMID:11016950
A;Accession: D84306
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-477 <STO>
A;Cross-references: GB:AE004437; NID:g10581016; PIDN:AAG19816.1; GSPDB:GN00138
C;Genetics:
A;Gene: nac
C;Superfamily: gamma-aminobutyric acid transporter

  Query Match             0.9%;  Score 8;  DB 2;  Length 477;
  Best Local Similarity   100.0%;  Pred. No. 34;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        358 AAVGSAVG 365
              ||||||||
Db         15 AAVGSAVG 22


RESULT 40
G84079
sodium-dependent transporter BH3439 [imported] - Bacillus halodurans (strain C-125)
C;Species: Bacillus halodurans
C;Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
C;Accession: G84079
R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; Hiran
Nucleic Acids Res. 28, 4317-4331, 2000
A;Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and g
A;Reference number: A83650; MUID:20512582; PMID:11058132
A;Accession: G84079
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-505 <STO>
A;Cross-references: GB:AP001518; GB:BA000004; NID:g10175792; PIDN:BAB07158.1; GSPDB:GN001
A;Experimental source: strain C-125
C;Genetics:
A;Gene: BH3439
C;Superfamily: gamma-aminobutyric acid transporter

  Query Match             0.9%;  Score 8;  DB 2;  Length 505;
  Best Local Similarity   100.0%;  Pred. No. 36;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        358 AAVGSAVG 365
              ||||||||
Db         14 AAVGSAVG 21


Search completed: February 17, 2004, 10:21:54
Job time : 57 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: February 17, 2004, 10:13:55 ; Search time 17 Seconds
(without alignments)
2564.339 Million cell updates/sec

Title: US-10-069-799-5
Perfect score: 927
Sequence: 1 MSNINVIKSNIQAGLNSTKS..........SSNALQPITQPTQGILAPSV 927

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 127863 seqs, 47026705 residues

Word size : 0

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : SwissProt_41:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	14	1.5	1023	1	HLY1_ECOLI	P09983 escherichia
2	14	1.5	1024	1	HLYA_ECOLI	P08715 escherichia
3	14	1.5	1049	1	RT31_ACTPL	P55130 actinobacil
4	14	1.5	1052	1	RT32_ACTPL	P55131 actinobacil
5	11	1.2	956	1	HLYA_ACTSU	Q00951 actinobacil
6	11	1.2	956	1	RT2A_ACTPL	P15377 actinobacil
7	11	1.2	1023	1	RT11_ACTPL	P55128 actinobacil
8	11	1.2	1023	1	RT12_ACTPL	P55129 actinobacil
9	10	1.1	947	1	LKTA_PASSP	P55123 pasteurella
10	10	1.1	953	1	LKA1_PASHA	P16535 pasteurella
11	10	1.1	953	1	LKA3_PASHA	P55116 pasteurella
12	10	1.1	953	1	LKAB_PASHA	P55118 pasteurella
13	10	1.1	955	1	LKAA_PASHA	P55117 pasteurella
14	8	0.9	262	1	YF14_MYCTU	P71793 mycobacteri
15	8	0.9	325	1	LIPA_CAUCR	Q9a7i8 caulobacter
16	8	0.9	433	1	Y384_MYCGE	P47624 mycoplasma
17	8	0.9	465	1	EGLC_RHIME	Q9z3q2 rhizobium m
18	8	0.9	773	1	MAK5_YEAST	P38112 saccharomyc
19	8	0.9	864	1	CHEA_BORBU	Q44737 borrelia bu
20	8	0.9	1050	1	LKTA_ACTAC	P16462 actinobacil
21	8	0.9	1705	1	CYAA_BORBR	Q57506 bordetella
22	7	0.8	18	1	RL24_PROVU	P20032 proteus vul
23	7	0.8	63	1	CECA_BOMMO	Q27239 bombyx mori
24	7	0.8	80	1	SCX1_TITSE	P01496 tityus serr
25	7	0.8	80	1	SCX4_TITSE	P45659 tityus serr
26	7	0.8	102	1	TRAC_RHISN	P55419 rhizobium s
27	7	0.8	103	1	RL24_ECOLI	P02425 escherichia
28	7	0.8	103	1	RL24_HAEIN	P44362 haemophilus
29	7	0.8	103	1	RL24_PASMU	Q9c141 pasteurella
30	7	0.8	104	1	RL24_BUCAK	P46177 buchnera ap
31	7	0.8	111	1	NU4M_CAICR	Q34076 caiman croc
32	7	0.8	122	1	RL7_NEILA	Q9etv2 neisseria l
33	7	0.8	122	1	RL7_NEIMA	P80716 neisseria m
34	7	0.8	122	1	RL7_YERPE	Q8zap4 yersinia pe
35	7	0.8	123	1	RL7_NEIPE	Q9f5m1 neisseria p
36	7	0.8	124	1	RL7_RALSO	Q8xuz7 ralstonia s
37	7	0.8	125	1	H2AV_STRPU	P08991 strongyloce
38	7	0.8	127	1	H2AV_CHICK	P02272 gallus gall
39	7	0.8	127	1	H2AZ_HUMAN	P17317 homo sapien
40	7	0.8	127	1	RL20_BIFLO	Q8g411 bifidobacte
41	7	0.8	131	1	BACH_HALAR	Q53461 haloarcula
42	7	0.8	140	1	H2AV_DROME	P08985 drosophila
43	7	0.8	141	1	MRAZ_MYCPN	P75467 mycoplasma
44	7	0.8	168	1	VTU2_DROME	P13238 drosophila
45	7	0.8	182	1	ATPF_RHORU	P15013 rhodospiril
46	7	0.8	186	1	NURM_NEUCR	P42116 neurospora
47	7	0.8	187	1	MT28_YEAST	P40573 saccharomyc
48	7	0.8	199	1	EFTS_THEMA	Q9x1u1 thermotoga
49	7	0.8	206	1	YIGZ_HAEIN	P44842 haemophilus
50	7	0.8	208	1	SXIG_VOLCA	P08471 volvox cart
51	7	0.8	210	1	AAC2_MYCSM	P94968 mycobacteri
52	7	0.8	215	1	CYB6_CYACA	Q9tlz7 cyanidium c
53	7	0.8	218	1	P25A_BOVIN	Q27957 bos taurus
54	7	0.8	219	1	P25A_HUMAN	O94811 homo sapien
55	7	0.8	220	1	MT04_BACHD	Q9k7u8 bacillus ha
56	7	0.8	223	1	GL1C_ARATH	Q9fma6 arabidopsis
57	7	0.8	226	1	COLI_ONCKE	P10000 o corticotr
58	7	0.8	229	1	RIBB_METTH	O27543 methanobact
59	7	0.8	232	1	TRPF_LIPST	Q01128 lipomyces s
60	7	0.8	235	1	AURB_CHLAU	P27197 chloroflexu
61	7	0.8	237	1	YCY1_SACER	P34698 saccharopol
62	7	0.8	240	1	COLJ_ONCMY	Q04618 oncorhynchu
63	7	0.8	254	1	GYLR_STRCO	P15360 streptomyce
64	7	0.8	256	1	CB4B_LYCES	P27525 lycopersico
65	7	0.8	256	1	NQRC_VIBCH	Q9x4q5 vibrio chol
66	7	0.8	260	1	COX3_XENLA	P00419 xenopus lae
67	7	0.8	260	1	REG2_DROME	Q94915 drosophila
68	7	0.8	261	1	COX3_BRARE	Q9miy4 brachydanio
69	7	0.8	261	1	COX3_GADMO	P55777 gadus morhu
70	7	0.8	261	1	COX3_POLOR	Q95914 polypterus
71	7	0.8	261	1	COX3_SQUAC	Q9zz48 squalus aca
72	7	0.8	275	1	PHAE_PHAVU	P05088 phaseolus v
73	7	0.8	276	1	BACH_HALHP	Q48315 halobacteri
74	7	0.8	276	1	BACH_HALHS	Q48314 halobacteri
75	7	0.8	279	1	SPBP_RAT	P08723 rattus norv
76	7	0.8	284	1	NODO_RHILV	P15728 rhizobium l
77	7	0.8	302	1	Y195_VIBCH	Q9kvf1 vibrio chol
78	7	0.8	307	1	RUVB_MYCGE	Q49425 mycoplasma
79	7	0.8	307	1	RUVB_MYCPN	P75242 mycoplasma
80	7	0.8	308	1	ERA_BRAJA	O69162 bradyrhizob
81	7	0.8	311	1	FLIM_BUCAP	Q8ka39 buchnera ap
82	7	0.8	320	1	ACC3_CUCME	P54847 cucumis mel
83	7	0.8	320	1	F4ST_FLACH	P52837 flaveria ch
84	7	0.8	323	1	CD47_HUMAN	Q08722 homo sapien
85	7	0.8	324	1	PSC2_MYCTU	O86344 mycobacteri
86	7	0.8	326	1	TGL2_YEAST	P54857 saccharomyc
87	7	0.8	331	1	PARG_HUMAN	Q9hbi0 homo sapien
88	7	0.8	340	1	RECA_RICPR	P41079 rickettsia
89	7	0.8	341	1	GPDA_LACLA	Q9cfx6 lactococcus
90	7	0.8	343	1	RECA_RICCN	Q92ge1 rickettsia
91	7	0.8	353	1	RRPO_PVSP	P22657 potato viru
92	7	0.8	360	1	VG47_BPMU	Q9t1v2 bacteriopha
93	7	0.8	363	1	OMPC_KLEPN	Q48473 klebsiella
94	7	0.8	366	1	AROC_NEIMA	Q9jt81 neisseria m
95	7	0.8	366	1	AROC_NEIMB	Q9jy99 neisseria m
96	7	0.8	370	1	AROG_CANAL	P79023 candida alb
97	7	0.8	372	1	AROC_XYLFA	Q9pd10 xylella fas
98	7	0.8	375	1	DNBI_HSVE1	Q03444 equine herp
99	7	0.8	376	1	OPPC_MYCPN	P75553 mycoplasma
100	7	0.8	380	1	CYB_LACVV	O48089 lacerta viv
101	7	0.8	380	1	PEL_PSEFL	Q59671 pseudomonas
102	7	0.8	380	1	TPSA_CAEEL	O77081 caenorhabdi
103	7	0.8	382	1	HMBP_DROME	P22809 drosophila
104	7	0.8	406	1	TRPB_CAUCR	P12290 caulobacter
105	7	0.8	414	1	GLYA_CAMJE	P24531 campylobact
106	7	0.8	448	1	G6PI_FUSNN	Q8rhh5 fusobacteri

107	7	0.8	449	1	MANB_METJA	Q57842	methanococc
108	7	0.8	454	1	NFM_PIG	P08552	sus scrofa
109	7	0.8	463	1	PGTP_SALTY	P12681	salmonella
110	7	0.8	463	1	YCOM_RHORU	P72315	rhodospiril
111	7	0.8	465	1	HNF6_HUMAN	Q9ubc0	homo sapien
112	7	0.8	465	1	HNF6_MOUSE	O08755	mus musculu
113	7	0.8	465	1	HNF6_RAT	P70512	rattus norv
114	7	0.8	467	1	FLII_BUCAI	P57178	buchnera ap
115	7	0.8	471	1	MURD_STRCO	Q9s2w9	streptomyce
116	7	0.8	475	1	PRTG_ERWCH	Q07162	erwinia chr
117	7	0.8	478	1	ASPA_ECOLI	P04422	escherichia
118	7	0.8	479	1	PRTC_ERWCH	P16317	erwinia chr
119	7	0.8	483	1	BCA_STRVL	P33569	streptomyce
120	7	0.8	486	1	PRZN_SERSP	P07268	serratia sp
121	7	0.8	487	1	PRZN_SERMA	P23694	serratia ma
122	7	0.8	492	1	SYK_THETH	P41255	thermus the
123	7	0.8	502	1	VL2_HPV48	Q80925	human papil
124	7	0.8	534	1	FMO2_CAVPO	P36366	cavia porce
125	7	0.8	537	1	CPA2_MUSDO	O18635	musca domes
126	7	0.8	538	1	NRM1_PIG	O77741	sus scrofa
127	7	0.8	547	1	GTR1_LEIDO	Q01440	leishmania
128	7	0.8	557	1	HNFB_HUMAN	P35680	homo sapien
129	7	0.8	557	1	HNFB_RAT	P23899	rattus norv
130	7	0.8	557	1	THS2_HALVO	O30560	halobacteri
131	7	0.8	558	1	HNFB_MOUSE	P27889	mus musculu
132	7	0.8	559	1	HNFB_PIG	Q03365	sus scrofa
133	7	0.8	560	1	TATR_NPVCF	P41716	choristoneu
134	7	0.8	585	1	YEJM_HAEIN	P44898	haemophilus
135	7	0.8	594	1	RMUC_NEIMA	Q9jwg3	neisseria m
136	7	0.8	594	1	RMUC_NEIMB	Q9jxh2	neisseria m
137	7	0.8	617	1	CYG2_HUMAN	O75343	homo sapien
138	7	0.8	622	1	S6A4_DROME	P51905	drosophila
139	7	0.8	635	1	SYT_METAC	Q8tlx7	methanosarc
140	7	0.8	637	1	DXS_RHILO	Q985y3	rhizobium l
141	7	0.8	639	1	DXS_AGRT5	Q8uhd7	agrobacteri
142	7	0.8	643	1	DXS_BRUME	Q8yfm2	brucella me
143	7	0.8	643	1	DXS_BRUSU	Q8g292	brucella su
144	7	0.8	645	1	DXS_RHIME	Q92rj1	rhizobium m
145	7	0.8	646	1	CG1I_CANAL	P24866	candida alb
146	7	0.8	646	1	T3MO_BPP1	P08763	bacteriopha
147	7	0.8	654	1	DNAK_CHLMU	P56836	chlamydia m
148	7	0.8	659	1	DNAK_CHLTR	P17821	chlamydia t
149	7	0.8	670	1	CATA_PENJA	P11934	penicillium
150	7	0.8	685	1	ATKB_CLOAB	O32328	clostridium
151	7	0.8	703	1	S261_RAT	P45380	rattus norv
152	7	0.8	781	1	COAT_PAVHB	P07299	human parvo
153	7	0.8	784	1	SX15_DROME	P40657	drosophila
154	7	0.8	792	1	Y85K_SSV1	P20210	sulfolobus
155	7	0.8	817	1	TGM1_HUMAN	P22735	homo sapien
156	7	0.8	827	1	PTF1_RHOCA	P23388	r multiphos
157	7	0.8	843	1	NUOG_STRCO	Q9xar0	streptomyce
158	7	0.8	847	1	DNLI_CAEEL	Q27474	caenorhabdi
159	7	0.8	986	1	RFC1_DROME	P35600	drosophila
160	7	0.8	1034	1	ACRF_ECOLI	P24181	escherichia
161	7	0.8	1041	1	TLR8_HUMAN	Q9nr97	homo sapien
162	7	0.8	1138	1	C7AA_BACTU	Q03749	bacillus th
163	7	0.8	1138	1	C7AB_BACUA	Q45707	bacillus th
164	7	0.8	1138	1	C7AB_BACUK	Q45708	bacillus th
165	7	0.8	1209	1	DNBI_HSVEB	P28932	equine herp
166	7	0.8	1287	1	VAC2_HELPY	Q48245	helicobacte
167	7	0.8	1288	1	VACA_HELPJ	Q9zkw5	helicobacte
168	7	0.8	1290	1	BXB_CLOBO	P10844	clostridium
169	7	0.8	1290	1	VACA_HELPY	P55981	helicobacte
170	7	0.8	1291	1	VAC4_HELPY	Q48258	helicobacte
171	7	0.8	1296	1	VAC1_HELPY	Q48247	helicobacte
172	7	0.8	1305	1	RPOC_UREPA	Q9pqv5	ureaplasma
173	7	0.8	1609	1	CTPI_MYCLE	O53114	mycobacteri
174	7	0.8	1625	1	CTPI_MYCTU	Q10900	mycobacteri
175	7	0.8	2300	1	CYAA_NEUCR	Q01631	neurospora
176	7	0.8	2314	1	AKA6_RAT	Q9wvc7	rattus norv
177	7	0.8	2319	1	AKA6_HUMAN	Q13023	homo sapien
178	7	0.8	2459	1	MAPB_RAT	P15205	rattus norv
179	7	0.8	2464	1	MAPB_MOUSE	P14873	mus musculu
180	7	0.8	2468	1	MAPB_HUMAN	P46821	homo sapien
181	7	0.8	2649	1	P285_HUMAN	Q9byk8	homo sapien
182	7	0.8	3135	1	S230_PLAFO	Q08372	plasmodium
183	7	0.8	3430	1	POLG_WNV	P06935	w genome po
184	7	0.8	3433	1	POLG_KUNJM	P14335	k genome po
185	7	0.8	3678	1	DMD_MOUSE	P11531	mus musculu
186	7	0.8	5147	1	PCLO_HUMAN	Q9y6v0	homo sapien
187	6	0.6	17	1	LPW_CORGL	P06556	corynebacte
188	6	0.6	31	1	PYSG_METBA	P80523	methanosarc
189	6	0.6	36	1	TLN1_CHICK	P54939	gallus gall
190	6	0.6	37	1	RL7_CLOPA	P05393	clostridium
191	6	0.6	38	1	RL7_VIBCO	P05395	vibrio cost
192	6	0.6	40	1	UC12_MAIZE	P80618	zea mays (m
193	6	0.6	42	1	MFM1_SCHPO	P34068	schizosacch
194	6	0.6	60	1	LHA2_RHOCA	P07367	rhodobacter
195	6	0.6	65	1	OMP_LOCMI	P80045	locusta mig
196	6	0.6	65	1	SAS2_BACCR	P06554	bacillus ce
197	6	0.6	71	1	IF1_LEPIN	Q9xd14	leptospira
198	6	0.6	71	1	YE34_ANASP	Q44148	anabaena sp
199	6	0.6	72	1	RPON_THEVO	Q979k0	thermoplasm
200	6	0.6	75	1	MCHB_ECOLI	Q9rm53	escherichia
201	6	0.6	77	1	DRG2_PHYBI	Q90zk5	phyllomedus
202	6	0.6	80	1	DMS3_PACDA	O93453	pachymedusa
203	6	0.6	81	1	DMS2_PHYBI	P31107	phyllomedus
204	6	0.6	83	1	MXIH_SHIFL	Q06079	shigella fl
205	6	0.6	84	1	EX7S_HAEIN	P43914	haemophilus
206	6	0.6	84	1	SASG_BACSU	P07784	bacillus su
207	6	0.6	84	1	UNGI_BPPB2	P14739	bacteriopha
208	6	0.6	84	1	YOAF_ECOLI	P76244	escherichia
209	6	0.6	85	1	RL28_FUSNN	Q8rdr9	fusobacteri
210	6	0.6	87	1	RPOH_THEVO	Q979f2	thermoplasm
211	6	0.6	94	1	CH10_THEBR	Q60023	thermoanaer
212	6	0.6	95	1	YD82_ARCFU	O28889	archaeoglob
213	6	0.6	95	1	YY1_ORYSA	O23810	oryza sativ
214	6	0.6	96	1	CH10_NEIMB	Q9jxm4	neisseria m
215	6	0.6	96	1	Y458_VIBCH	Q9kuq7	vibrio chol
216	6	0.6	98	1	IM08_SCHPO	Q09783	schizosacch
217	6	0.6	100	1	REGN_BPP22	P04891	bacteriopha
218	6	0.6	102	1	ACYM_ANAPL	P14620	anas platyr
219	6	0.6	102	1	ACYM_CHICK	P07031	gallus gall
220	6	0.6	102	1	ACYM_MELGA	P00821	meleagris g
221	6	0.6	103	1	C552_NITEU	P95339	nitrosomona
222	6	0.6	103	1	RL24_BACSU	P12876	bacillus su
223	6	0.6	103	1	RR6_CYACA	O19917	cyanidium c
224	6	0.6	104	1	YGY9_YEAST	P53069	saccharomyc
225	6	0.6	105	1	RL24_THEMA	P38513	thermotoga
226	6	0.6	105	1	Y3C3_STRCO	Q53867	streptomyce
227	6	0.6	106	1	Y0B5_MYCLE	Q49723	mycobacteri
228	6	0.6	106	1	YCIH_HAEIN	P45116	haemophilus
229	6	0.6	107	1	IHFA_BRUME	Q8ygi0	brucella me
230	6	0.6	107	1	YI46_SYNY3	P73056	synechocyst
231	6	0.6	108	1	HDEB_ECOLI	P26605	escherichia
232	6	0.6	109	1	Y268_CAUCR	Q9abf9	caulobacter
233	6	0.6	109	1	ZM33_MAIZE	O82106	zea mays (m
234	6	0.6	110	1	HIS2_CLOAB	Q97kh6	clostridium
235	6	0.6	110	1	LOL1_MOUSE	P97873	mus musculu
236	6	0.6	112	1	IHFA_AGRT5	Q8ug61	agrobacteri
237	6	0.6	112	1	IHFA_RHIME	Q92qt3	rhizobium m
238	6	0.6	114	1	SMS1_ICTPU	P01171	ictalurus p
239	6	0.6	115	1	PT19_STYPL	P28211	styela plic
240	6	0.6	116	1	RL18_MYCCA	P04453	mycoplasma
241	6	0.6	117	1	RL20_CAMJE	Q9piq0	campylobact
242	6	0.6	118	1	RL20_AQUAE	O67086	aquifex aeo
243	6	0.6	118	1	YE16_HAEIN	P44188	haemophilus
244	6	0.6	118	1	YI5C_ECOLI	P19770	escherichia
245	6	0.6	120	1	NU3C_LUPLU	P52765	lupinus lut
246	6	0.6	120	1	RL7_ECOLI	P02392	escherichia
247	6	0.6	120	1	RL7_SALTY	P18081	salmonella
248	6	0.6	121	1	RL7_CLOPE	Q8xhr7	clostridium
249	6	0.6	121	1	RL7_UREPA	Q9prd4	ureaplasma
250	6	0.6	121	1	RL7_VIBCH	Q9kv31	vibrio chol
251	6	0.6	121	1	RL7_XANAC	Q8pnt1	xanthomonas
252	6	0.6	121	1	RL7_XANCP	Q8pc57	xanthomonas

253	6	0.6	122	1	RL7_HAEIN	P44348	haemophilus
254	6	0.6	122	1	RL7_HALEU	P07472	halophilic
255	6	0.6	122	1	RL7_HALPR	P14134	haloanaerob
256	6	0.6	122	1	RL7_PASMU	Q9ck90	pasteurella
257	6	0.6	122	1	RL7_VIBPA	Q87kq3	vibrio para
258	6	0.6	122	1	RL7_XYLFA	Q9pa85	xylella fas
259	6	0.6	122	1	YMI9_YEAST	Q04502	saccharomyc
260	6	0.6	123	1	RL7_CLOAB	Q97eg8	clostridium
261	6	0.6	123	1	VST1_HEVBU	P29325	hepatitis e
262	6	0.6	123	1	VST1_HEVME	Q03499	hepatitis e
263	6	0.6	123	1	VST1_HEVMY	Q04612	hepatitis e
264	6	0.6	124	1	ACPS_CLOAB	Q97lr5	clostridium
265	6	0.6	124	1	R22B_ARATH	Q9m9w1	arabidopsis
266	6	0.6	124	1	R22C_ARATH	Q9fe58	arabidopsis
267	6	0.6	124	1	RL7_BRUME	P41106	brucella me
268	6	0.6	125	1	C556_AGRTA	P00141	agrobacteri
269	6	0.6	125	1	RL7_AGRT5	Q8ue07	agrobacteri
270	6	0.6	125	1	RL7_LIBAF	P41189	liberibacte
271	6	0.6	125	1	RL7_RHILO	Q98n67	rhizobium l
272	6	0.6	125	1	RL7_RICCN	Q92j87	rickettsia
273	6	0.6	125	1	RL7_RICPR	Q9ze21	rickettsia
274	6	0.6	125	1	RL7_THETN	Q8r7u5	thermoanaer
275	6	0.6	125	1	YOM2_PHOPR	P29740	photobacter
276	6	0.6	126	1	RL7_DESVM	P02393	desulfovibr
277	6	0.6	126	1	RL7_RHIME	Q92qh8	rhizobium m
278	6	0.6	127	1	KRCL_CHICK	P25692	gallus gall
279	6	0.6	127	1	R22A_ARATH	Q9srx7	arabidopsis
280	6	0.6	127	1	RL7_CAUCR	Q9aau7	caulobacter
281	6	0.6	127	1	RL7_SYNY3	P23349	synechocyst
282	6	0.6	127	1	RT08_ACACA	P46757	acanthamoeb
283	6	0.6	128	1	GCSH_YERPE	Q8zhi7	yersinia pe
284	6	0.6	128	1	KDGL_HELPJ	Q9zle0	helicobacte
285	6	0.6	128	1	KDGL_HELPY	P56411	helicobacte
286	6	0.6	128	1	RL7_AQUAE	O67761	aquifex aeo
287	6	0.6	129	1	LV1D_MOUSE	P01726	mus musculu
288	6	0.6	129	1	MFA4_BOVIN	P55918	bos taurus
289	6	0.6	129	1	RL7_ANASP	Q8ylj5	anabaena sp
290	6	0.6	129	1	TDCF_ECOLI	P42631	escherichia
291	6	0.6	129	1	YB58_METJA	Q58557	methanococc
292	6	0.6	130	1	F14A_HUMAN	Q9h2x8	homo sapien
293	6	0.6	130	1	RL14_LEPIN	Q9xd26	leptospira
294	6	0.6	130	1	Y122_BUCAI	P57222	buchnera ap
295	6	0.6	130	1	Y483_DEIRA	Q9rx32	deinococcus
296	6	0.6	133	1	GOP_BPP4	P13058	bacteriopha
297	6	0.6	134	1	H2AV_YEAST	Q12692	saccharomyc
298	6	0.6	134	1	Y886_HAEIN	P44069	haemophilus
299	6	0.6	135	1	LGUL_HAEIN	P44638	haemophilus
300	6	0.6	136	1	YGHG_ECOLI	Q46835	escherichia
301	6	0.6	137	1	GTH2_ACALA	Q90225	acanthopagr
302	6	0.6	137	1	RK19_NEPOL	Q9tkz4	nephroselmi
303	6	0.6	138	1	CXA4_PIG	Q29559	sus scrofa
304	6	0.6	139	1	YQGF_HAEIN	P43981	haemophilus
305	6	0.6	140	1	HBB_RANES	P02134	rana escule
306	6	0.6	140	1	POP7_HUMAN	O75817	homo sapien
307	6	0.6	141	1	HBA1_CATMA	P82111	catharacta
308	6	0.6	141	1	HBA1_IGUIG	P18974	iguana igua
309	6	0.6	141	1	HBA2_BUBBU	Q9tsn8	bubalus bub
310	6	0.6	141	1	HBAD_ANAPL	P04442	anas platyr
311	6	0.6	141	1	HBAD_ANSAN	P04238	anser anser
312	6	0.6	141	1	HBAD_ANSIN	P04239	anser indic
313	6	0.6	141	1	HBAD_BRACA	P04240	branta cana
314	6	0.6	141	1	HBAD_CAIMO	P02003	cairina mos
315	6	0.6	141	1	HBAD_CHLME	P07035	chloephaga
316	6	0.6	141	1	HBA_ANAPE	P22740	anas penelo
317	6	0.6	141	1	HBA_ANAPL	P01986	anas platyr
318	6	0.6	141	1	HBA_ANAPP	P01988	anas platyr
319	6	0.6	141	1	HBA_ANSAN	P01989	anser anser
320	6	0.6	141	1	HBA_ANSSE	P01985	anseranas s
321	6	0.6	141	1	HBA_APTFO	P01980	aptenodytes
322	6	0.6	141	1	HBA_ARAAR	P01996	ara araraun
323	6	0.6	141	1	HBA_BRACA	P01991	branta cana
324	6	0.6	141	1	HBA_CAIMO	P01987	cairina mos
325	6	0.6	141	1	HBA_CHLME	P07034	chloephaga
326	6	0.6	141	1	HBA_COTJA	P24589	coturnix co
327	6	0.6	141	1	HBA_CYGOL	P01992	cygnus olor
328	6	0.6	141	1	HBA_LARRI	P08260	larus ridib
329	6	0.6	141	1	HBA_PASMO	P07407	passer mont
330	6	0.6	141	1	HBA_PHACA	P10780	phalacrocor
331	6	0.6	141	1	HBA_RAT	P01946	rattus norv
332	6	0.6	141	1	HBA_STUVU	P01997	sturnus vul
333	6	0.6	141	1	HBA_TURME	P14522	turdus meru
334	6	0.6	141	1	PRO1_STRPU	P32006	strongyloce
335	6	0.6	141	1	RL23_TRYCR	Q94776	trypanosoma
336	6	0.6	141	1	YQBS_BACSU	P45934	bacillus su
337	6	0.6	142	1	YQAA_ECOLI	P76631	escherichia
338	6	0.6	143	1	VG32_VZVD	P09285	varicella-z
339	6	0.6	143	1	YIBN_ECOLI	P37688	escherichia
340	6	0.6	144	1	FLAV_TRIER	O52659	trichodesmi
341	6	0.6	145	1	PA21_BOVIN	P00593	bos taurus
342	6	0.6	146	1	RPB8_YEAST	P20436	saccharomyc
343	6	0.6	146	1	YRFB_ECOLI	P45751	escherichia
344	6	0.6	147	1	YA37_MYCPN	P75077	mycoplasma
345	6	0.6	148	1	CTR6_SCHPO	Q9usv7	schizosacch
346	6	0.6	148	1	PSAL_SYNEN	P25902	synechococc
347	6	0.6	149	1	IL2_HORSE	P37997	equus cabal
348	6	0.6	149	1	MLE_BRAFL	Q17133	branchiosto
349	6	0.6	149	1	YB29_MYCPN	P75346	mycoplasma
350	6	0.6	150	1	EXBB_HAEDU	O51808	haemophilus
351	6	0.6	150	1	MLE_DICDI	P09402	dictyosteli
352	6	0.6	150	1	NIFW_RHISN	P55689	rhizobium s
353	6	0.6	150	1	PA2C_MOUSE	P48076	mus musculu
354	6	0.6	150	1	PA2C_RAT	P39878	rattus norv
355	6	0.6	150	1	PSAL_ODOSI	P49486	odontella s
356	6	0.6	151	1	YZ28_AQUAE	O66419	aquifex aeo
357	6	0.6	152	1	EXBB_PASHA	P72202	pasteurella
358	6	0.6	152	1	IAA2_HORVU	P13691	hordeum vul
359	6	0.6	152	1	PYRI_THEVO	Q97b28	thermoplasm
360	6	0.6	153	1	IL2_HUMAN	P01585	homo sapien
361	6	0.6	153	1	YIBG_ECOLI	P32106	escherichia
362	6	0.6	154	1	IL2_CERTO	P46649	cercocebus
363	6	0.6	154	1	IL2_FELCA	Q07885	felis silve
364	6	0.6	154	1	IL2_MACFA	Q29615	macaca fasc
365	6	0.6	154	1	IL2_MACMU	P51498	macaca mula
366	6	0.6	154	1	IL2_MIRAN	O62641	mirounga an
367	6	0.6	154	1	YQAR_BACSU	P45914	bacillus su
368	6	0.6	155	1	Y115_METJA	Q57579	methanococc
369	6	0.6	156	1	ATPF_VIBAL	P12989	vibrio algi
370	6	0.6	156	1	FMAD_BACNO	P13253	bacteroides
371	6	0.6	156	1	VG55_BPML5	Q05265	mycobacteri
372	6	0.6	157	1	EGD1_YEAST	Q02642	saccharomyc
373	6	0.6	157	1	FKB1_METJA	Q57726	methanococc
374	6	0.6	157	1	NIFU_AQUAE	O67045	aquifex aeo
375	6	0.6	157	1	PR1_MEDSA	Q43560	medicago sa
376	6	0.6	157	1	SSSP_PLAFO	P17503	plasmodium
377	6	0.6	157	1	YG17_PASMU	Q9ckj8	pasteurella
378	6	0.6	158	1	DRR3_PEA	P14710	pisum sativ
379	6	0.6	158	1	RK12_ORYSA	O22386	oryza sativ
380	6	0.6	159	1	RL22_MYCPN	P75575	mycoplasma
381	6	0.6	159	1	YIB6_YEAST	P40548	saccharomyc
382	6	0.6	160	1	BCHF_RHOSH	Q53222	rhodobacter
383	6	0.6	160	1	PHAA_CYAPA	P00316	cyanophora
384	6	0.6	161	1	7B2_XENLA	P18844	xenopus lae
385	6	0.6	161	1	PHAB_AGLNE	P28556	aglaothamni
386	6	0.6	162	1	DYR_LACCA	P00381	lactobacill
387	6	0.6	162	1	MOTB_BPT4	Q01437	bacteriopha
388	6	0.6	163	1	ATPX_OCHNE	Q40608	ochrosphaer
389	6	0.6	163	1	GCSH_MESCR	P93255	mesembryant
390	6	0.6	163	1	MLC2_YEAST	Q06580	saccharomyc
391	6	0.6	164	1	YAF9_METAC	P58828	methanosarc
392	6	0.6	165	1	OGT_BACSU	P11742	bacillus su
393	6	0.6	166	1	PMPA_CANBO	P14292	candida boi
394	6	0.6	166	1	PMPB_CANBO	P14293	candida boi
395	6	0.6	166	1	RL10_PSEAE	Q9hwc7	pseudomonas
396	6	0.6	166	1	RS5_STAAM	Q99s38	staphylococ
397	6	0.6	167	1	DYR_ENTFC	P00380	enterococcu
398	6	0.6	167	1	RS5_SHEON	P59124	shewanella

399	6	0.6	169	1	YP03_BORBU	Q44849 borrelia bu
400	6	0.6	170	1	LED1_METKA	Q8tx94 methanopyru
401	6	0.6	171	1	CYP7_CAEEL	P52015 caenorhabdi
402	6	0.6	171	1	RL10_CORGL	Q8nt29 corynebacte
403	6	0.6	172	1	CYP1_ARATH	P34790 arabidopsis
404	6	0.6	172	1	CYP2_CAEEL	P52010 caenorhabdi
405	6	0.6	173	1	CYP3_CAEEL	P52011 caenorhabdi
406	6	0.6	173	1	VB06_VACCV	P21002 vaccinia vi
407	6	0.6	173	1	Y096_NPVAC	P41656 autographa
408	6	0.6	174	1	AXI5_ARATH	P33078 arabidopsis
409	6	0.6	174	1	HSCB_HAEIN	Q57006 haemophilus
410	6	0.6	175	1	CYT2_STOHE	P07845 stoichactis
411	6	0.6	175	1	ET2_MOUSE	P22389 mus musculu
412	6	0.6	175	1	NU6M_CANFA	Q9zz56 canis famil
413	6	0.6	175	1	YF49_MYCTU	Q10777 mycobacteri
414	6	0.6	176	1	CYT1_STOHE	P81662 stoichactis
415	6	0.6	176	1	PACA_HUMAN	P18509 h pituitary
416	6	0.6	178	1	ARGR_STRCL	P95721 streptomyce
417	6	0.6	178	1	ATPD_STRBO	O50156 streptococc
418	6	0.6	179	1	ARGR_STRCO	Q911a5 streptomyce
419	6	0.6	179	1	Y053_SYNY3	Q55142 synechocyst
420	6	0.6	180	1	PORC_METTH	O27772 methanobact
421	6	0.6	181	1	IF3_RHOSH	O33567 rhodobacter
422	6	0.6	181	1	Y06E_BPT4	P13311 bacteriopha
423	6	0.6	182	1	APT_CAMJE	Q9pp06 campylobact
424	6	0.6	182	1	RELX_PIG	P01348 sus scrofa
425	6	0.6	183	1	YGIL_ECOLI	P39834 escherichia
426	6	0.6	184	1	RK5_MESVI	Q9muu5 mesostigma
427	6	0.6	184	1	YCF4_CHAGL	Q8m9x4 chaetosphae
428	6	0.6	185	1	PFD3_DROME	Q9vgp6 drosophila
429	6	0.6	185	1	RRF_VIBPA	Q8grf5 vibrio para
430	6	0.6	186	1	RPO6_FOWPV	Q85280 fowlpox vir
431	6	0.6	186	1	YB15_ARCFU	O29150 archaeoglob
432	6	0.6	187	1	KAD_XANAC	Q8ph23 xanthomonas
433	6	0.6	188	1	HPPK_MYCTU	O06276 mycobacteri
434	6	0.6	188	1	KTHY_METJA	Q57741 methanococc
435	6	0.6	188	1	UBCB_YEAST	P52491 saccharomyc
436	6	0.6	189	1	NUXM_NEUCR	Q02854 neurospora
437	6	0.6	189	1	RMP2_MOUSE	Q9wup0 mus musculu
438	6	0.6	189	1	Y573_METJA	Q57993 methanococc
439	6	0.6	191	1	KADA_METKA	Q8tzb0 methanopyru
440	6	0.6	191	1	KIP1_HUMAN	Q99828 homo sapien
441	6	0.6	191	1	KIP1_MOUSE	Q9z0f4 mus musculu
442	6	0.6	191	1	KIP1_RAT	Q9r010 rattus norv
443	6	0.6	192	1	CYP1_CAEEL	P52009 caenorhabdi
444	6	0.6	194	1	Y583_BUCAP	O51880 buchnera ap
445	6	0.6	194	1	YZ10_AQUAE	O66404 aquifex aeo
446	6	0.6	195	1	KAD_SYNEL	Q8dml4 synechococc
447	6	0.6	195	1	TRPG_CYACA	O19914 cyanidium c
448	6	0.6	196	1	PINQ_ECOLI	P77170 escherichia
449	6	0.6	196	1	PINR_ECOLI	P77574 escherichia
450	6	0.6	196	1	RNH2_AQUAE	O67768 aquifex aeo
451	6	0.6	196	1	RS5_DEIRA	Q9rs11 deinococcus
452	6	0.6	198	1	RACD_ENTHI	Q24817 entamoeba h
453	6	0.6	199	1	YE65_MYCPN	P75318 mycoplasma
454	6	0.6	200	1	COAE_BRUME	Q8ye21 brucella me
455	6	0.6	200	1	RBFA_ARATH	O65693 arabidopsis
456	6	0.6	200	1	SYST_LYCES	P27058 lycopersico
457	6	0.6	201	1	PSS_METJA	Q58609 methanococc
458	6	0.6	201	1	RECR_ECOLI	P12727 escherichia
459	6	0.6	201	1	RECR_SALTY	Q8xg25 salmonella
460	6	0.6	201	1	RUVA_CHLTE	Q8kfr0 chlorobium
461	6	0.6	202	1	CO8G_RABIT	Q28679 oryctolagus
462	6	0.6	202	1	GDIR_YEAST	Q12434 saccharomyc
463	6	0.6	204	1	MSX3_MOUSE	P70354 mus musculu
464	6	0.6	204	1	YA57_ACTAC	O52728 actinobacil
465	6	0.6	205	1	3MGH_CLOPE	Q8xha9 clostridium
466	6	0.6	205	1	RUVA_AGRT5	Q8u9k5 agrobacteri
467	6	0.6	205	1	YD06_AGRT5	Q8uful agrobacteri
468	6	0.6	206	1	RR4_ODOSI	P49492 odontella s
469	6	0.6	207	1	CYPB_CHICK	P24367 gallus gall
470	6	0.6	207	1	YJY2_YEAST	P47087 saccharomyc
471	6	0.6	210	1	CD8B_FELCA	P79336 felis silve
472	6	0.6	210	1	CD8B_HUMAN	P10966 homo sapien
473	6	0.6	210	1	CD8B_PONPY	P30434 pongo pygma
474	6	0.6	210	1	KTHY_PYRFU	Q8u071 pyrococcus
475	6	0.6	211	1	HET3_RADMG	Q9u6x1 radianthus
476	6	0.6	211	1	ROPB_RHILV	Q52866 rhizobium l
477	6	0.6	211	1	SDF2_HUMAN	Q99470 homo sapien
478	6	0.6	211	1	SDF2_MOUSE	Q9dct5 mus musculu
479	6	0.6	211	1	YH73_METTH	O27801 methanobact
480	6	0.6	212	1	IL6_CERTO	P46650 cercocebus
481	6	0.6	212	1	IL6_HUMAN	P05231 homo sapien
482	6	0.6	212	1	IL6_MACFA	P79341 macaca fasc
483	6	0.6	212	1	IL6_MACMU	P51494 macaca mula
484	6	0.6	212	1	PCAJ_PSEPK	Q01104 pseudomonas
485	6	0.6	212	1	RPSD_MYCTU	Q50712 mycobacteri
486	6	0.6	213	1	HIS5_LEGPN	Q9rdx3 legionella
487	6	0.6	213	1	LEXA_XANAC	O86050 xanthomonas
488	6	0.6	213	1	LEXA_XANCA	Q93mq9 xanthomonas
489	6	0.6	213	1	LEXA_XANCP	Q8p9x2 xanthomonas
490	6	0.6	213	1	PURQ_THEMA	Q9x0x2 thermotoga
491	6	0.6	213	1	PYRE_HAEIN	P43855 haemophilus
492	6	0.6	213	1	Y767_MYCTU	P71822 mycobacteri
493	6	0.6	213	1	YE54_HAEIN	P44202 haemophilus
494	6	0.6	213	1	YJH2_YEAST	P40359 saccharomyc
495	6	0.6	214	1	GIDB_PSEAE	Q9ht10 pseudomonas
496	6	0.6	215	1	Y315_TREPA	P56822 treponema p
497	6	0.6	217	1	PSBP_BRAJU	Q96334 brassica ju
498	6	0.6	217	1	SWSH_DROME	Q9vj26 drosophila
499	6	0.6	219	1	NEP1_SULTO	Q96yp4 sulfolobus
500	6	0.6	219	1	RAS_SCHPO	P08647 schizosacch
501	6	0.6	219	1	RT02_ARATH	Q9gcb9 arabidopsis
502	6	0.6	219	1	YGHB_SALTY	P18951 salmonella
503	6	0.6	220	1	CLD6_HUMAN	P56747 homo sapien
504	6	0.6	220	1	RADB_PYRKO	P95547 pyrococcus
505	6	0.6	221	1	GTA1_CHICK	Q08392 gallus gall
506	6	0.6	221	1	SDFL_HUMAN	Q9hcn8 homo sapien
507	6	0.6	221	1	SDFL_MOUSE	Q9esp1 mus musculu
508	6	0.6	221	1	TCPP_VIBCH	P29485 vibrio chol
509	6	0.6	222	1	COLI_CYPCA	Q9ygk4 c corticotr
510	6	0.6	222	1	KCY_AQUAE	O67907 aquifex aeo
511	6	0.6	222	1	SODM_HORSE	Q9xs41 equus cabal
512	6	0.6	223	1	PLLV_RAT	P34207 rattus norv
513	6	0.6	223	1	RPIA_STRA3	Q8e4y6 streptococc
514	6	0.6	223	1	RPIA_STRA5	Q8dzc6 streptococc
515	6	0.6	224	1	MTGA_ACICA	O24849 acinetobact
516	6	0.6	224	1	MYOG_MOUSE	P12979 mus musculu
517	6	0.6	224	1	PURQ_HALN1	Q9hnu2 halobacteri
518	6	0.6	226	1	MODB_AZOVI	P37731 azotobacter
519	6	0.6	227	1	KTHY_XANAC	Q8pfg7 xanthomonas
520	6	0.6	228	1	CLDA_HUMAN	P78369 homo sapien
521	6	0.6	229	1	ISTB_BURCE	P55923 burkholderi
522	6	0.6	229	1	PYRH_THEAC	Q9hkv7 thermoplasm
523	6	0.6	229	1	Y4MB_RHISN	P55561 rhizobium s
524	6	0.6	229	1	Y756_AQUAE	O66958 aquifex aeo
525	6	0.6	229	1	YO14_ARCFU	O30257 archaeoglob
526	6	0.6	230	1	C552_THIFE	P74917 thiobacillu
527	6	0.6	230	1	YGEA_ECOLI	P03813 escherichia
528	6	0.6	231	1	CLDA_MOUSE	Q9z0s6 mus musculu
529	6	0.6	232	1	CLP2_ANASP	Q8yqx8 anabaena sp
530	6	0.6	232	1	UPP_PYRAB	Q9v0k1 pyrococcus
531	6	0.6	232	1	UPP_PYRFU	Q8u1g7 pyrococcus
532	6	0.6	233	1	MTGA_NEIGO	Q51005 neisseria g
533	6	0.6	233	1	MTGA_NEIMA	O52423 neisseria m
534	6	0.6	233	1	VRPP_LAMBD	P03689 bacteriopha
535	6	0.6	233	1	Y081_NPVAC	Q06694 autographa
536	6	0.6	234	1	YX02_CAEEL	Q11109 caenorhabdi
537	6	0.6	235	1	TFE2_ALCEU	P94136 alcaligenes
538	6	0.6	235	1	Y381_MYCPN	P75219 mycoplasma
539	6	0.6	235	1	Y680_RHILO	Q98m91 rhizobium l
540	6	0.6	236	1	PCNA_SOYBN	P22177 glycine max
541	6	0.6	237	1	NOLA_BRASN	P50330 bradyrhizob
542	6	0.6	237	1	PHOU_ZYMMO	Q9x5e1 zymomonas m
543	6	0.6	237	1	PYRF_FUSNN	Q8rg83 fusobacteri
544	6	0.6	237	1	SFSA_SYNY3	P73664 synechocyst

545	6	0.6	238	1	Y939_HAEIN	P44080	haemophilus
546	6	0.6	239	1	RS3_YEAST	P05750	saccharomyc
547	6	0.6	240	1	DAPB_STAAM	Q99u88	staphylococ
548	6	0.6	240	1	DAPB_STAAU	Q9ez11	staphylococ
549	6	0.6	240	1	DAPB_STAAW	Q8nws4	staphylococ
550	6	0.6	240	1	RR40_YEAST	Q08285	saccharomyc
551	6	0.6	241	1	AGL8_SINAL	Q41274	sinapis alb
552	6	0.6	241	1	FIMC_ECOL6	P59590	escherichia
553	6	0.6	241	1	FIMC_ECOLI	P31697	escherichia
554	6	0.6	241	1	FLGF_RHIME	O54248	rhizobium m
555	6	0.6	241	1	NAGB_CLOAB	Q97mk9	clostridium
556	6	0.6	241	1	PSB1_YEAST	P23724	saccharomyc
557	6	0.6	242	1	AGL8_ARATH	Q38876	arabidopsis
558	6	0.6	242	1	CD8A_BOVIN	P31783	bos taurus
559	6	0.6	242	1	HEM3_CHLMU	Q9pk95	chlamydia m
560	6	0.6	242	1	MPGP_PYRFU	Q8u381	pyrococcus
561	6	0.6	242	1	NAGB_CLOPE	Q8xhp8	clostridium
562	6	0.6	243	1	YO27_CAEEL	P34677	caenorhabdi
563	6	0.6	245	1	PDXJ_SHEON	Q8eh78	shewanella
564	6	0.6	246	1	MCT9_MOUSE	O35164	mus musculu
565	6	0.6	246	1	MCTX_MOUSE	Q00356	mus musculu
566	6	0.6	247	1	GRAB_MOUSE	P04187	mus musculu
567	6	0.6	248	1	CUTC_ECOLI	P46719	escherichia
568	6	0.6	248	1	GRL1_RAT	Q06605	rattus norv
569	6	0.6	248	1	NKP1_RAT	P18291	rattus norv
570	6	0.6	248	1	PDXJ_PSEAE	Q9i5g5	pseudomonas
571	6	0.6	248	1	RPOE_SULAC	P39466	sulfolobus
572	6	0.6	248	1	TRPA_ARCFU	O28673	archaeoglob
573	6	0.6	248	1	YT31_RALSO	Q8xva0	ralstonia s
574	6	0.6	249	1	RADC_BRUME	Q8yhs8	brucella me
575	6	0.6	251	1	LP4_BOMMO	P09337	bombyx mori
576	6	0.6	251	1	Y545_METJA	Q57965	methanococc
577	6	0.6	251	1	YJFQ_ECOLI	P39299	escherichia
578	6	0.6	252	1	C4BB_HUMAN	P20851	homo sapien
579	6	0.6	252	1	PMG2_HUMAN	P15259	homo sapien
580	6	0.6	252	1	TRT3_COTJA	P06398	coturnix co
581	6	0.6	253	1	3BHD_COMTE	P19871	comamonas t
582	6	0.6	253	1	TPM1_HYDAT	P39921	hydra atten
583	6	0.6	253	1	TRUA_LACLA	Q9ci80	lactococcus
584	6	0.6	253	1	VG23_HSVSA	Q01006	herpesvirus
585	6	0.6	254	1	GLVR_BACSU	P54717	bacillus su
586	6	0.6	254	1	PHSC_SALTY	P37602	salmonella
587	6	0.6	254	1	Y351_CAUCR	Q9ab80	caulobacter
588	6	0.6	254	1	YH11_MYCTU	O33210	mycobacteri
589	6	0.6	254	1	YMH1_CAEEL	P34468	caenorhabdi
590	6	0.6	255	1	1433_ARATH	P42644	arabidopsis
591	6	0.6	255	1	HIS9_CLOAB	P58291	clostridium
592	6	0.6	255	1	MFA4_HUMAN	P55083	homo sapien
593	6	0.6	255	1	OP32_DROME	Q9vcg3	drosophila
594	6	0.6	256	1	6PGL_CHLTR	O84189	chlamydia t
595	6	0.6	256	1	BIOH_ECOLI	P13001	escherichia
596	6	0.6	256	1	MYF5_NOTVI	Q91154	notophthalm
597	6	0.6	257	1	KLAA_ECOLI	Q57239	escherichia
598	6	0.6	257	1	Y125_PYRHO	O57865	pyrococcus
599	6	0.6	257	1	Y271_BUCAP	P42395	buchnera ap
600	6	0.6	258	1	HB21_SPAEH	P15464	spalax leuc
601	6	0.6	258	1	MCR_SHEEP	Q9bdj7	ovis aries
602	6	0.6	258	1	RFA2_CRIFA	Q23697	crithidia f
603	6	0.6	258	1	TPIS_STRCO	Q9z520	streptomyce
604	6	0.6	259	1	DHSB_PARDE	Q59662	paracoccus
605	6	0.6	259	1	PSBP_PEA	P16059	pisum sativ
606	6	0.6	259	1	YTMN_BACSU	O34900	bacillus su
607	6	0.6	260	1	PSBP_SINAL	P11594	sinapis alb
608	6	0.6	260	1	VBR1_ICMV	Q08595	indian cass
609	6	0.6	260	1	VLT1_VACCV	P21029	vaccinia vi
610	6	0.6	260	1	VLT1_VARV	P32990	variola vir
611	6	0.6	261	1	1438_LYCES	P93213	lycopersico
612	6	0.6	261	1	RPOD_PYRFU	Q8u0e4	pyrococcus
613	6	0.6	261	1	RS2_PEDAC	P49668	pediococcus
614	6	0.6	262	1	TRT3_CHICK	P12620	gallus gall
615	6	0.6	263	1	CB23_ORYSA	P27519	oryza sativ
616	6	0.6	263	1	L302_BOMMO	Q00801	bombyx mori
617	6	0.6	263	1	PCN1_ARATH	Q9m7q7	arabidopsis
618	6	0.6	263	1	PCNA_BRANA	Q43124	brassica na
619	6	0.6	263	1	PCNA_MAIZE	Q43266	zea mays (m
620	6	0.6	263	1	PCNA_ORYSA	P17070	oryza sativ
621	6	0.6	263	1	PSP1_ARATH	Q42029	arabidopsis
622	6	0.6	264	1	GLPF_HAEIN	P44826	haemophilus
623	6	0.6	264	1	PCN2_ARATH	Q9zw35	arabidopsis
624	6	0.6	264	1	PCNA_DAUCA	Q00268	daucus caro
625	6	0.6	264	1	PCNA_POPNI	Q9may3	populus nig
626	6	0.6	264	1	PCNA_STYCL	P53358	styela clav
627	6	0.6	264	1	PCNA_TOBAC	O82797	nicotiana t
628	6	0.6	264	1	UPK_BUCAP	Q8ka52	buchnera ap
629	6	0.6	265	1	H1_PEA	P08283	pisum sativ
630	6	0.6	265	1	HXDB_HETFR	Q9ia15	heterodontu
631	6	0.6	265	1	PSBP_TOBAC	P18212	nicotiana t
632	6	0.6	265	1	UPK_CYTJO	Q9rb37	cytophaga j
633	6	0.6	266	1	NADE_METTH	O27554	methanobact
634	6	0.6	266	1	NANH_BACFR	P31206	bacteroides
635	6	0.6	266	1	THID_RHIME	P56904	rhizobium m
636	6	0.6	267	1	PARA_CAUCR	O05189	caulobacter
637	6	0.6	267	1	PSBP_SPIOL	P12302	spinacia ol
638	6	0.6	267	1	STO3_CAEEL	Q20657	caenorhabdi
639	6	0.6	268	1	MIND_BACSU	Q01464	bacillus su
640	6	0.6	268	1	PCNA_CATRO	P24314	catharanthu
641	6	0.6	268	1	TRPA_PSEAE	P07344	pseudomonas
642	6	0.6	269	1	AAC3_SERMA	Q01515	serratia ma
643	6	0.6	269	1	DAPB_NEIMA	Q9jx48	neisseria m
644	6	0.6	269	1	DAPB_NEIMB	Q9k1f1	neisseria m
645	6	0.6	269	1	MIND_ECOLI	P18197	escherichia
646	6	0.6	269	1	RECX_LISIN	Q92aw7	listeria in
647	6	0.6	269	1	TRPA_PSEPU	P11081	pseudomonas
648	6	0.6	269	1	TRPA_PSESY	P34816	pseudomonas
649	6	0.6	270	1	DAPB_HAEIN	P45153	haemophilus
650	6	0.6	270	1	Y619_TREPA	O83628	treponema p
651	6	0.6	271	1	ALLR_ECOLI	P77734	escherichia
652	6	0.6	271	1	HXA9_MOUSE	P09631	mus musculu
653	6	0.6	271	1	MIND_THEMA	Q9x2i3	thermotoga
654	6	0.6	271	1	STM1_SCHPO	Q10482	schizosacch
655	6	0.6	271	1	YW34_MYCTU	O05879	mycobacteri
656	6	0.6	273	1	DAB1_RHILO	P58210	rhizobium l
657	6	0.6	273	1	FLJL_CAUCR	P18914	caulobacter
658	6	0.6	273	1	H12_ARATH	P26569	arabidopsis
659	6	0.6	273	1	MURI_SHEON	Q8ek90	shewanella
660	6	0.6	273	1	NADC_HELPJ	Q9zjn2	helicobacte
661	6	0.6	273	1	NADC_HELPY	O25909	helicobacte
662	6	0.6	273	1	Y133_METJA	Q57597	methanococc
663	6	0.6	274	1	BAG1_HUMAN	Q99933	homo sapien
664	6	0.6	274	1	OMPZ_CHLPN	Q9z6i6	chlamydia p
665	6	0.6	274	1	PCNA_PLAFK	P31008	plasmodium
666	6	0.6	274	1	PDX4_BOVIN	Q9bgi2	bos taurus
667	6	0.6	274	1	RHAD_ECOLI	P32169	escherichia
668	6	0.6	274	1	YL14_ARCFU	O28166	archaeoglob
669	6	0.6	275	1	APAH_ACTAC	O52655	actinobacil
670	6	0.6	275	1	APAH_PASMU	P57922	pasteurella
671	6	0.6	275	1	HPCB_ECOLI	Q05353	escherichia
672	6	0.6	275	1	TRPA_CAUCR	P12291	caulobacter
673	6	0.6	275	1	TSG6_MOUSE	O08859	mus musculu
674	6	0.6	275	1	Y05B_BPT4	P39257	bacteriopha
675	6	0.6	276	1	TSG6_RABIT	P98065	oryctolagus
676	6	0.6	277	1	ICE3_CRILO	Q60431	cricetulus
677	6	0.6	277	1	ICE3_HUMAN	P42574	homo sapien
678	6	0.6	277	1	ICE3_MOUSE	P70677	mus musculu
679	6	0.6	277	1	ICE3_RAT	P55213	rattus norv
680	6	0.6	277	1	RK2_SPIMX	O98452	spirogyra m
681	6	0.6	277	1	TSG6_HUMAN	P98066	homo sapien
682	6	0.6	277	1	UBS1_YEAST	P38290	saccharomyc
683	6	0.6	278	1	ERA_BUCAP	Q8k9r2	buchnera ap
684	6	0.6	278	1	NIFH_METVO	P06119	methanococc
685	6	0.6	278	1	T2D7_DROME	Q27272	drosophila
686	6	0.6	278	1	YEIG_ECOLI	P33018	escherichia
687	6	0.6	279	1	GLPF_PSEAE	Q51389	pseudomonas
688	6	0.6	279	1	LGTB_NEIGO	Q50947	neisseria g
689	6	0.6	279	1	LGTB_NEIMA	P57033	neisseria m
690	6	0.6	279	1	NADE_SULTO	Q96yl5	sulfolobus

691	6	0.6	279	1	NIFH_METJA	Q58289	methanococc
692	6	0.6	279	1	YB49_STRP3	Q8k6t3	streptococc
693	6	0.6	279	1	YE93_STRPY	Q99yy0	streptococc
694	6	0.6	280	1	ISPE_OCEIH	Q8eu37	oceanobacil
695	6	0.6	280	1	NUDC_DEIRA	Q9rv62	deinococcus
696	6	0.6	280	1	PANC_XANAC	Q8p110	xanthomonas
697	6	0.6	280	1	PURU_ECOL6	P38480	escherichia
698	6	0.6	280	1	PURU_ECOLI	P37051	escherichia
699	6	0.6	280	1	YO17_MYCTU	P71726	mycobacteri
700	6	0.6	282	1	EFTS_CHLPN	Q9z7k8	chlamydia p
701	6	0.6	282	1	FDHD_THEAC	Q9h117	thermoplasm
702	6	0.6	282	1	Y125_THEMA	Q9wxx9	thermotoga
703	6	0.6	283	1	CG1C_CHICK	P55168	gallus gall
704	6	0.6	283	1	DAPF_NEIMA	Q9jv69	neisseria m
705	6	0.6	283	1	DAPF_NEIMB	Q9k060	neisseria m
706	6	0.6	283	1	YQVW_CAEEL	Q17439	caenorhabdi
707	6	0.6	284	1	FOCA_HAEIN	P43756	haemophilus
708	6	0.6	284	1	PANC_XANCP	Q8p9s9	xanthomonas
709	6	0.6	285	1	AROE_CAUCR	Q9ac57	caulobacter
710	6	0.6	285	1	ARSB_MOUSE	P50429	mus musculu
711	6	0.6	285	1	PANC_CAUCR	Q9a6c8	caulobacter
712	6	0.6	286	1	PPNK_CAMJE	Q9phm6	campylobact
713	6	0.6	286	1	TRT2_SHEEP	P50751	ovis aries
714	6	0.6	286	1	YDBC_ECOLI	P25906	escherichia
715	6	0.6	287	1	FOLD_SALTI	Q60006	salmonella
716	6	0.6	287	1	FOLD_SALTY	P58688	salmonella
717	6	0.6	287	1	HBD_BACSU	P45856	bacillus su
718	6	0.6	287	1	TYSY_MYCGE	P47469	mycoplasma
719	6	0.6	288	1	CD80_HUMAN	P33681	homo sapien
720	6	0.6	288	1	OTX2_XENLA	Q91813	xenopus lae
721	6	0.6	288	1	TRA6_NEIMB	Q00840	neisseria m
722	6	0.6	289	1	OTX2_BRARE	Q91981	brachydanio
723	6	0.6	289	1	OTX2_HUMAN	P32243	homo sapien
724	6	0.6	289	1	OTX2_MOUSE	P80206	mus musculu
725	6	0.6	290	1	BLA1_XANMA	P52700	xanthomonas
726	6	0.6	290	1	YC38_CYAPA	P48278	cyanophora
727	6	0.6	291	1	ISPA_MICLU	O66126	micrococcus
728	6	0.6	291	1	KPRS_METKA	Q8tut6	methanopyru
729	6	0.6	292	1	ARGB_METMA	Q8pxj8	methanosarc
730	6	0.6	293	1	PRMA_ECOLI	P28637	escherichia
731	6	0.6	294	1	LEP_MYCTU	Q10789	mycobacteri
732	6	0.6	294	1	YHDJ_ECOLI	P28638	escherichia
733	6	0.6	295	1	ISPE_PASMU	P57833	pasteurella
734	6	0.6	296	1	MY88_HUMAN	Q99836	homo sapien
735	6	0.6	296	1	YC32_METJA	Q58629	methanococc
736	6	0.6	297	1	BPHC_BURCE	P47228	burkholderi
737	6	0.6	297	1	HIS1_YEAST	P00498	saccharomyc
738	6	0.6	297	1	ISPE_LEPIN	Q8ezm8	leptospira
739	6	0.6	297	1	PECM_ERWCH	P42194	erwinia chr
740	6	0.6	297	1	Y060_MYCGE	P47306	mycoplasma
741	6	0.6	298	1	CG1C_RAT	P39947	rattus norv
742	6	0.6	298	1	YHAJ_ECOLI	P42623	escherichia
743	6	0.6	298	1	YIHU_ECOLI	P32142	escherichia
744	6	0.6	299	1	ISPE_AGRT5	Q8uhp8	agrobacteri
745	6	0.6	299	1	ISPE_STRCO	Q9k3r6	streptomyce
746	6	0.6	299	1	SCRK_BACSU	O05510	bacillus su
747	6	0.6	299	1	YF02_MYCTU	P71783	mycobacteri
748	6	0.6	300	1	Y921_METJA	Q58331	methanococc
749	6	0.6	301	1	FPG_RHIME	Q59752	rhizobium m
750	6	0.6	301	1	GCVA_HAEIN	P45099	haemophilus
751	6	0.6	301	1	HMGL_PSEMV	P13703	pseudomonas
752	6	0.6	301	1	PYRD_THETN	Q8r9r7	thermoanaer
753	6	0.6	302	1	BPHC_PSEPS	P08695	pseudomonas
754	6	0.6	302	1	PP12_DROME	P12982	drosophila
755	6	0.6	302	1	PP13_DROME	Q05547	drosophila
756	6	0.6	302	1	RBCR_CHRVI	P25544	chromatium
757	6	0.6	303	1	CG1C_HUMAN	P24863	homo sapien
758	6	0.6	303	1	MIAA_SYNY3	P74040	synechocyst
759	6	0.6	303	1	T2RD_HUMAN	Q9nyv9	homo sapien
760	6	0.6	304	1	CG1C_MOUSE	Q62447	mus musculu
761	6	0.6	304	1	SUAC_RAT	P50237	rattus norv
762	6	0.6	305	1	ARCC_THEVO	Q97c45	thermoplasm
763	6	0.6	305	1	ERA_PSEAE	Q9xcx8	pseudomonas
764	6	0.6	305	1	GP7D_CHLMU	Q46435	chlamydia m
765	6	0.6	305	1	GP7D_CHLTR	P10561	chlamydia t
766	6	0.6	305	1	PYRB_SERMA	P19910	serratia ma
767	6	0.6	305	1	RS1B_SYNY3	P74142	synechocyst
768	6	0.6	305	1	SGAU_MYCPN	P75294	mycoplasma
769	6	0.6	305	1	YYAM_BACSU	P37511	bacillus su
770	6	0.6	306	1	ARGI_BRUME	Q59174	brucella me
771	6	0.6	306	1	UCP1_MOUSE	P12242	mus musculu
772	6	0.6	306	1	UCP1_RAT	P04633	rattus norv
773	6	0.6	307	1	CIW7_HUMAN	Q9y2u2	homo sapien
774	6	0.6	307	1	GP7D_CHLPS	Q46257	chlamydia p
775	6	0.6	307	1	PRB2_CORGL	Q8ns12	corynebacte
776	6	0.6	307	1	YHB3_YEAST	P38747	saccharomyc
777	6	0.6	308	1	ACP1_ENTHI	P36184	entamoeba h
778	6	0.6	308	1	AEF1_DROME	P39413	drosophila
779	6	0.6	308	1	CYB_AMBMA	P29632	amblyornis
780	6	0.6	308	1	CYB_AMPST	P29633	ampelion st
781	6	0.6	308	1	CYB_COLRU	P29635	colaptes ru
782	6	0.6	308	1	CYB_PARIN	P29638	parus inorn
783	6	0.6	308	1	CYB_PTIPL	P29640	ptiloprora
784	6	0.6	308	1	CYB_SCYMA	P29641	scytalopus
785	6	0.6	308	1	CYB_THRDO	P29642	thripophaga
786	6	0.6	309	1	FABH_AQUAE	O67185	aquifex aeo
787	6	0.6	309	1	NTCB_SYNP7	P52693	synechococc
788	6	0.6	309	1	PYRB_VIBPA	Q871f8	vibrio para
789	6	0.6	309	1	PYRB_VIBS2	P96174	vibrio sp.
790	6	0.6	309	1	PYRB_VIBVU	Q8dcf6	vibrio vuln
791	6	0.6	310	1	PYRB_ECOLI	P00479	escherichia
792	6	0.6	310	1	PYRB_SALTY	P08420	salmonella
793	6	0.6	310	1	RPOA_CYACA	Q9tlv2	cyanidium c
794	6	0.6	310	1	YCJY_ECOLI	P76049	escherichia
795	6	0.6	311	1	F3ST_FLACH	P52836	flaveria ch
796	6	0.6	311	1	PYRB_YERPE	Q8zb39	yersinia pe
797	6	0.6	311	1	YXXF_BACSU	Q07835	bacillus su
798	6	0.6	312	1	F3ST_FLABI	P52835	flaveria bi
799	6	0.6	312	1	FABD_HAEIN	P43712	haemophilus
800	6	0.6	312	1	GUB2_HORVU	P12257	hordeum vul
801	6	0.6	312	1	O1F1_HUMAN	O43749	homo sapien
802	6	0.6	312	1	OXYR_STRVD	Q9x5p2	streptomyce
803	6	0.6	312	1	RPOA_LACLA	Q9cdy3	lactococcus
804	6	0.6	312	1	T2R6_MOUSE	P59530	mus musculu
805	6	0.6	313	1	OYE6_HUMAN	Q96rd3	homo sapien
806	6	0.6	313	1	YGCG_ECOLI	P55140	escherichia
807	6	0.6	313	1	YO06_YEAST	Q12094	saccharomyc
808	6	0.6	314	1	BLA2_STRCI	P14560	streptomyce
809	6	0.6	314	1	FMT_BUCAI	P57564	buchnera ap
810	6	0.6	314	1	HCDH_HUMAN	Q16836	homo sapien
811	6	0.6	314	1	SERB_HAEIN	P44997	haemophilus
812	6	0.6	315	1	YZ09_AQUAE	O66403	aquifex aeo
813	6	0.6	316	1	ARGI_XENLA	P30759	xenopus lae
814	6	0.6	316	1	KHSE_MYCTU	Q10603	mycobacteri
815	6	0.6	316	1	MRAW_VIBCH	Q9kpf9	vibrio chol
816	6	0.6	316	1	MRAW_VIBPR	Q9ajg9	vibrio prot
817	6	0.6	316	1	MRAW_VIBVU	Q8dek2	vibrio vuln
818	6	0.6	316	1	PECT_ERWCH	P52662	erwinia chr
819	6	0.6	316	1	TNP1_HUMAN	Q13829	homo sapien
820	6	0.6	316	1	YNF8_YEAST	P53947	saccharomyc
821	6	0.6	317	1	FMT_BACSU	P94463	bacillus su
822	6	0.6	317	1	FMT_FUSNN	Q8rdm3	fusobacteri
823	6	0.6	317	1	PSRC_WOLSU	P31077	wolinella s
824	6	0.6	317	1	RL0A_ARATH	O04204	arabidopsis
825	6	0.6	318	1	CLH2_ARATH	Q9m7i7	arabidopsis
826	6	0.6	318	1	CYF_CHAGL	Q8m9x2	chaetosphae
827	6	0.6	318	1	OYD1_HUMAN	Q9h346	homo sapien
828	6	0.6	318	1	T2R7_HUMAN	Q9nyw3	homo sapien
829	6	0.6	318	1	YCCG_BACSU	P70954	bacillus su
830	6	0.6	319	1	HEMZ_STRMU	Q8cww4	streptococc
831	6	0.6	319	1	PTE1_HUMAN	O14734	homo sapien
832	6	0.6	319	1	RLA0_ORYSA	P41095	oryza sativ
833	6	0.6	320	1	ASPG_SPOFR	O02467	spodoptera
834	6	0.6	320	1	PTE1_MOUSE	P58137	mus musculu
835	6	0.6	321	1	EAEB_ECO27	Q05129	escherichia
836	6	0.6	321	1	IF2B_SCHPO	P56329	schizosacch

837	6	0.6	321	1	RBSC_ECOLI	P04984 escherichia
838	6	0.6	321	1	SFX3_HUMAN	Q9bwm7 homo sapien
839	6	0.6	321	1	SFX3_MOUSE	Q91v61 mus musculu
840	6	0.6	321	1	SFX3_RAT	Q9jhy2 rattus norv
841	6	0.6	322	1	ARGI_HUMAN	P05089 homo sapien
842	6	0.6	322	1	GRP2_MOUSE	O89100 m grb2-rela
843	6	0.6	322	1	MIAA_WIGBR	Q8d318 wiggleswort
844	6	0.6	322	1	NADA_CHLTE	Q8kew2 chlorobium
845	6	0.6	322	1	POR_SYNY3	Q59987 synechocyst
846	6	0.6	322	1	RLA0_LUPLU	P50345 lupinus lut
847	6	0.6	322	1	SFX1_HUMAN	Q9h9b4 homo sapien
848	6	0.6	322	1	SFX1_MOUSE	Q99jr1 mus musculu
849	6	0.6	322	1	SFX1_RAT	Q63965 rattus norv
850	6	0.6	322	1	YQJA_BACSU	P54538 bacillus su
851	6	0.6	323	1	ARGI_MOUSE	Q61176 mus musculu
852	6	0.6	323	1	ARGI_RAT	P07824 rattus norv
853	6	0.6	323	1	HEMZ_HAEIN	P43868 haemophilus
854	6	0.6	323	1	VP9_RGDV	P23628 rice gall d
855	6	0.6	324	1	CAHF_MOUSE	Q99n23 mus musculu
856	6	0.6	324	1	ODPB_BACSU	P21882 bacillus su
857	6	0.6	324	1	YL70_ARCFU	O28112 archaeoglob
858	6	0.6	325	1	ACC2_DORSP	Q39705 doritaenops
859	6	0.6	325	1	LACD_STRMU	P26425 streptococc
860	6	0.6	325	1	PSBO_WHEAT	P27665 triticum ae
861	6	0.6	325	1	RPOA_MYCGA	Q9rdv6 mycoplasma
862	6	0.6	325	1	Y132_CHLPN	Q9z949 chlamydia p
863	6	0.6	326	1	HEMZ_PASMU	P57874 pasteurella
864	6	0.6	326	1	LACD_LACLA	P26593 lactococcus
865	6	0.6	326	1	LACD_STAAM	P11100 staphylococ
866	6	0.6	326	1	NU1M_CHOCR	P48898 chondrus cr
867	6	0.6	326	1	VS09_ROTEL	P25176 equine rota
868	6	0.6	326	1	VS09_ROTPM	P32549 porcine rot
869	6	0.6	327	1	CCG4_HUMAN	Q9ubn1 homo sapien
870	6	0.6	327	1	CCG4_MOUSE	Q9jjv4 mus musculu
871	6	0.6	327	1	CCG4_RAT	Q8vhw9 rattus norv
872	6	0.6	327	1	FKB6_HUMAN	O75344 homo sapien
873	6	0.6	327	1	ODPB_ACHLA	P35488 acholeplasm
874	6	0.6	327	1	PP11_DROME	P48461 drosophila
875	6	0.6	327	1	PSTC_RHILO	Q98f13 rhizobium l
876	6	0.6	327	1	TAL_CHLMU	Q9pk80 chlamydia m
877	6	0.6	327	1	TAL_CHLTR	O84315 chlamydia t
878	6	0.6	328	1	HAM1_STRP3	Q8k8i7 streptococc
879	6	0.6	328	1	HAM1_STRP8	Q8p2d2 streptococc
880	6	0.6	328	1	HAM1_STRPY	Q9a1b6 streptococc
881	6	0.6	328	1	I12B_HUMAN	P29460 homo sapien
882	6	0.6	328	1	I12B_MACMU	P48095 macaca mula
883	6	0.6	328	1	K6PF_MYCPN	P75476 mycoplasma
884	6	0.6	328	1	NAS6_HORVU	Q9zqv3 hordeum vul
885	6	0.6	328	1	NODZ_AZOCA	Q43966 azorhizobiu
886	6	0.6	328	1	SYFA_BUCBP	P59504 buchnera ap
887	6	0.6	328	1	YJ77_YEAST	P47145 saccharomyc
888	6	0.6	328	1	YRBH_ECOLI	P45395 escherichia
889	6	0.6	329	1	CAHX_FLAPR	P46281 flaveria pr
890	6	0.6	329	1	IDI2_STRPY	Q9a095 streptococc
891	6	0.6	329	1	NAS4_HORVU	Q9zqv6 hordeum vul
892	6	0.6	329	1	NAS7_HORVU	Q9zwh8 hordeum vul
893	6	0.6	329	1	NAS8_HORVU	Q9xfb6 hordeum vul
894	6	0.6	329	1	PDXA_NEIMA	Q9jx42 neisseria m
895	6	0.6	329	1	PDXA_NEIMB	Q9k1f9 neisseria m
896	6	0.6	329	1	PSBO_FRIAG	O49079 fritillaria
897	6	0.6	329	1	PSBO_LYCES	P23322 lycopersico
898	6	0.6	329	1	PSBO_PEA	P14226 pisum sativ
899	6	0.6	329	1	SRPG_SYNP7	Q59966 synechococc
900	6	0.6	329	1	Y461_CHLTR	O84467 chlamydia t
901	6	0.6	329	1	YCLC_PSEPU	Q47100 pseudomonas
902	6	0.6	330	1	CAH1_FLALI	P46512 flaveria li
903	6	0.6	330	1	CAHX_FLABR	P46511 flaveria br
904	6	0.6	330	1	GRP2_HUMAN	O75791 h grb2-rela
905	6	0.6	330	1	PEXD_CAEEL	Q19951 caenorhabdi
906	6	0.6	330	1	RPOS_ECOLI	P13445 escherichia
907	6	0.6	330	1	RPOS_SALTY	P37400 salmonella
908	6	0.6	330	1	RPOS_SHIFL	P35540 shigella fl
909	6	0.6	331	1	CAHX_FLABI	P46510 flaveria bi
910	6	0.6	331	1	MACS_BOVIN	P12624 bos taurus
911	6	0.6	331	1	PSO2_ARATH	Q9s841 arabidopsis
912	6	0.6	331	1	RL3_ARCFU	O28354 archaeoglob
913	6	0.6	331	1	RPOS_YEREN	P47765 yersinia en
914	6	0.6	331	1	YB30_ARATH	O82314 arabidopsis
915	6	0.6	332	1	HRDD_STRCO	P18249 streptomyce
916	6	0.6	332	1	HRDD_STRVF	Q60012 streptomyce
917	6	0.6	332	1	HS30_YEAST	P25619 saccharomyc
918	6	0.6	332	1	PSBO_TOBAC	Q40459 nicotiana t
919	6	0.6	332	1	PSO1_ARATH	P23321 arabidopsis
920	6	0.6	333	1	PSBO_SOLTU	P26320 solanum tub
921	6	0.6	333	1	VINT_BPMD2	Q38361 mycobacteri
922	6	0.6	333	1	VINT_BPMFR	P25426 mycobacteri
923	6	0.6	334	1	CATL_MOUSE	P06797 mus musculu
924	6	0.6	334	1	CATL_RAT	P07154 rattus norv
925	6	0.6	334	1	UNG_HSV11	P10186 herpes simp
926	6	0.6	334	1	Y472_RICPR	Q9zd72 rickettsia
927	6	0.6	335	1	ILC1_SULSO	Q9uwx9 sulfolobus
928	6	0.6	335	1	OTCC_PSEAE	P08308 pseudomonas
929	6	0.6	335	1	RPOS_VIBCH	O51804 vibrio chol
930	6	0.6	335	1	STS1_ARATH	P94111 arabidopsis
931	6	0.6	336	1	ARGC_MORS3	Q9k4z1 moritella s
932	6	0.6	336	1	ASNA_CLOPE	Q8xiy5 clostridium
933	6	0.6	336	1	IDI2_STRPN	Q97sh8 streptococc
934	6	0.6	336	1	LEU3_SULTO	P50455 sulfolobus
935	6	0.6	336	1	MRAY_STRP3	Q8k6c7 streptococc
936	6	0.6	336	1	MRAY_STRP8	Q8nzy2 streptococc
937	6	0.6	336	1	MRAY_STRPY	Q99yk2 streptococc
938	6	0.6	337	1	RPOA_ORYSA	P12090 oryza sativ
939	6	0.6	337	1	TRPD_BRAJA	P94326 bradyrhizob
940	6	0.6	338	1	YDJ4_SCHPO	P87052 schizosacch
941	6	0.6	339	1	LPXD_YERPE	P58611 yersinia pe
942	6	0.6	339	1	PAX9_CHICK	P55166 gallus gall
943	6	0.6	339	1	RT13_YEAST	P12686 saccharomyc
944	6	0.6	339	1	YE60_XYLFA	Q9pdb9 xylella fas
945	6	0.6	340	1	THPS_SULAC	P17118 sulfolobus
946	6	0.6	340	1	YM21_ARCFU	O28062 archaeoglob
947	6	0.6	341	1	MDHM_CAEEL	O02640 caenorhabdi
948	6	0.6	341	1	MRAW_RHIME	Q92n14 rhizobium m
949	6	0.6	341	1	OMPU_VIBCH	P97085 vibrio chol
950	6	0.6	341	1	POLG_HPAV1	P13672 hepatitis a
951	6	0.6	342	1	CITC_KLEPN	P45410 klebsiella
952	6	0.6	342	1	LYCA_BPCP7	P19385 bacteriopha
953	6	0.6	342	1	NMT1_ASPPA	P42882 aspergillus
954	6	0.6	342	1	TRM1_THEVO	Q97ar2 thermoplasm
955	6	0.6	343	1	HRCA_BACSU	P25499 bacillus su
956	6	0.6	343	1	SRT1_YEAST	Q03175 saccharomyc
957	6	0.6	343	1	VIBY_AGRTU	P05360 agrobacteri
958	6	0.6	344	1	CGG2_MOUSE	O08918 mus musculu
959	6	0.6	344	1	COA2_POVJC	P03095 polyomaviru
960	6	0.6	344	1	FLIM_TREPA	P74927 treponema p
961	6	0.6	344	1	NU2M_RANCA	P16673 rana catesb
962	6	0.6	344	1	SYW_CHLPN	Q9z7a4 chlamydia p
963	6	0.6	345	1	D2D2_XENLA	P34973 xenopus lae
964	6	0.6	346	1	CDK7_HUMAN	P50613 homo sapien
965	6	0.6	346	1	EPR1_AERHY	O05485 aeromonas h
966	6	0.6	346	1	GALM_ECOLI	P40681 escherichia
967	6	0.6	346	1	PSTS_ECOLI	P06128 escherichia
968	6	0.6	346	1	SYW_CHLMU	Q9pjf5 chlamydia m
969	6	0.6	346	1	SYW_CHLTR	O84589 chlamydia t
970	6	0.6	346	1	TAS_ECOLI	Q46933 escherichia
971	6	0.6	346	1	YE77_METJA	Q58872 methanococc
972	6	0.6	347	1	ITP2_HUMAN	Q9ukp3 homo sapien
973	6	0.6	347	1	NADA_ECO57	Q8x963 escherichia
974	6	0.6	347	1	NADA_ECOLI	P11458 escherichia
975	6	0.6	347	1	NADA_SALTI	Q8z8b8 salmonella
976	6	0.6	347	1	NADA_SALTY	P24519 salmonella
977	6	0.6	347	1	NADA_SHIFL	Q93cp9 shigella fl
978	6	0.6	348	1	STRI_STRGR	P09400 streptomyce
979	6	0.6	349	1	ESR1_ANOCA	Q9yht3 anolis caro
980	6	0.6	349	1	LPXD_ANASP	Q8ys10 anabaena sp
981	6	0.6	350	1	MRP_DEIRA	Q9rvm9 deinococcus
982	6	0.6	350	1	V0D2_DROME	Q9vcq3 drosophila

```
  983     6   0.6    350  1  YEEL_ECOLI          P76349 escherichia
  984     6   0.6    351  1  COA2_POVBA          P14997 polyomaviru
  985     6   0.6    351  1  COA2_POVBK          P03094 polyomaviru
  986     6   0.6    351  1  FTSY_MYCHO          O32861 mycoplasma
  987     6   0.6    351  1  LPXK_XANCP          Q8p8w5 xanthomonas
  988     6   0.6    352  1  DP42_SULSO          Q97w02 sulfolobus
  989     6   0.6    352  1  FLIM_BORBU          Q57511 borrelia bu
  990     6   0.6    352  1  HEMZ_PROFR          P72183 propionibac
  991     6   0.6    352  1  LPCC_RHILV          O68547 rhizobium l
  992     6   0.6    352  1  WN4A_BRARE          P47793 brachydanio
  993     6   0.6    353  1  ADA_SALTY           P26189 salmonella
  994     6   0.6    353  1  CEBA_BOVIN          O02754 bos taurus
  995     6   0.6    353  1  LPXD_CHLTE          Q8kcq3 chlorobium
  996     6   0.6    353  1  NADA_YERPE          Q8zgy8 yersinia pe
  997     6   0.6    353  1  YRRI_BACSU          O34472 bacillus su
  998     6   0.6    354  1  YCJF_HAEIN          P43931 haemophilus
  999     6   0.6    355  1  BAG1_MOUSE          Q60739 mus musculu
 1000     6   0.6    355  1  BUK_LISMO           Q8y7b6 listeria mo
```

ALIGNMENTS

```
RESULT 1
HLY1_ECOLI
ID   HLY1_ECOLI     STANDARD;      PRT;  1023 AA.
AC   P09983;
DT   01-MAR-1989 (Rel. 10, Created)
DT   01-MAR-1989 (Rel. 10, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Hemolysin, chromosomal.
GN   HLYA.
OS   Escherichia coli.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Escherichia.
OX   NCBI_TaxID=562;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=J96 / Serotype O4;
RX   MEDLINE=85234404; PubMed=3891743;
RA   Felmlee T., Pellett S., Welch R.A.;
RT   "Nucleotide sequence of an Escherichia coli chromosomal hemolysin.";
RL   J. Bacteriol. 163:94-105(1985).
RN   [2]
RP   SEQUENCE OF 1-44 FROM N.A.
RC   STRAIN=2001;
RX   MEDLINE=85258115; PubMed=3894051;
RA   Nicaud J.-M., Mackman N., Gray L., Holland I.B.;
RT   "Characterisation of HlyC and mechanism of activation and secretion
RT   of haemolysin from E. coli 2001.";
RL   FEBS Lett. 187:339-344(1985).
CC   -!- FUNCTION: Bacterial hemolysins are exotoxins that attack blood
CC       cell membranes and cause cell rupture by mechanisms not clearly
CC       defined.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- DOMAIN: The Gly-rich region is probably involved in binding
CC       calcium, which is required for target cell-binding or cytolytic
CC       activity.
CC   -!- DOMAIN: The three transmembrane domains are believed to be
CC       involved in pore formation by the cytotoxin.
CC   -!- PTM: Palmitoylated by hlyC. The toxin only becomes active when
CC       modified.
CC   -!- MISCELLANEOUS: The hemolysin of E.coli is produced predominantly
CC       by strains causing extraintestinal infections, such as those of
CC       the urinary tract.
CC   -!- SIMILARITY: BELONGS TO THE RTX PROKARYOTIC TOXIN FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; M10133; AAA23975.1; -.
DR   EMBL; X02768; CAA26546.1; -.
DR   PIR; A24433; LEECA.
DR   InterPro; IPR001343; Hemlysn_Ca_bind.
DR   InterPro; IPR003995; RtxA.
DR   Pfam; PF00353; hemolysinCabind; 6.
DR   Pfam; PF02382; RTX; 1.
DR   PRINTS; PR00313; CABNDNGRPT.
DR   PRINTS; PR01488; RTXTOXINA.
DR   PROSITE; PS00330; HEMOLYSIN_CALCIUM; 4.
KW   Hemolysis; Toxin; Cytolysis; Repeat; Calcium; Transmembrane;
KW   Lipoprotein; Palmitate.
FT   TRANSMEM    237    259       POTENTIAL.
FT   TRANSMEM    267    326       POTENTIAL.
FT   TRANSMEM    364    410       POTENTIAL.
FT   DOMAIN      723    869       16 X REPEATS, GLY-RICH.
FT   REPEAT      723    728       1.
FT   REPEAT      732    737       2.
FT   REPEAT      741    746       3.
FT   REPEAT      750    755       4.
FT   REPEAT      759    764       5.
FT   REPEAT      768    773       6.
FT   REPEAT      777    782       7.
FT   REPEAT      786    791       8.
FT   REPEAT      795    800       9.
FT   REPEAT      806    812       10.
FT   REPEAT      816    821       11.
FT   REPEAT      825    830       12.
FT   REPEAT      834    839       13.
FT   REPEAT      843    848       14.
FT   REPEAT      855    860       15.
FT   REPEAT      864    869       16.
FT   LIPID       563    563       PALMITATE (BY SIMILARITY).
FT   LIPID       689    689       PALMITATE (BY SIMILARITY).
FT   VARIANT       6      6       A -> T (IN STRAIN 2001).
SQ   SEQUENCE   1023 AA;  109867 MW;  196D5C0A9A28B54D CRC64;

  Query Match             1.5%;  Score 14;  DB 1;  Length 1023;
  Best Local Similarity   100.0%;  Pred. No. 2.2e-05;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        379 ISGILEASKQAMFE 392
              ||||||||||||||
Db        405 ISGILEASKQAMFE 418


RESULT 2
HLYA_ECOLI
ID   HLYA_ECOLI     STANDARD;      PRT;  1024 AA.
AC   P08715;
DT   01-JAN-1988 (Rel. 06, Created)
DT   01-JAN-1988 (Rel. 06, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Hemolysin, plasmid.
GN   HLYA.
OS   Escherichia coli.
OG   Plasmid IncI2 pHLY152.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Escherichia.
OX   NCBI_TaxID=562;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Hess J., Wels W., Vogel M., Goebel W.;
RT   "Nucleotide sequence of a plasmid-encoded hemolysin determinant and
RT   its comparison with a corresponding chromosomal hemolysin sequence.";
RL   FEMS Microbiol. Lett. 34:1-11(1986).
RN   [2]
RP   PALMITOYLATION OF LYS-564 AND LYS-690.
RX   MEDLINE=95099325; PubMed=7801126;
RA   Stanley P., Packman L.C., Koronakis V., Hughes C.;
RT   "Fatty acylation of two internal lysine residues required for the
```

```
RT   toxic activity of Escherichia coli hemolysin.";
RL   Science 266:1992-1996(1994).
RN   [3]
RP   PALMITOYLATION OF LYS-564 AND LYS-690.
RX   MEDLINE=96404790; PubMed=8808931;
RA   Ludwig A., Garcia F., Bauer S., Jarchau T., Benz R., Hoppe J.,
RA   Goebel W.;
RT   "Analysis of the in vivo activation of hemolysin (HlyA) from
RT   Escherichia coli.";
RL   J. Bacteriol. 178:5422-5430(1996).
CC   -!- FUNCTION: Bacterial hemolysins are exotoxins that attack blood
CC       cell membranes and cause cell rupture by mechanisms not clearly
CC       defined.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- DOMAIN: The Gly-rich region is probably involved in binding
CC       calcium, which is required for target cell-binding or cytolytic
CC       activity.
CC   -!- DOMAIN: The three transmembrane domains are believed to be
CC       involved in pore formation by the cytotoxin.
CC   -!- PTM: Palmitoylated by hlyC. The toxin only becomes active when
CC       modified.
CC   -!- MISCELLANEOUS: The hemolysin of E.Coli is produced predominantly
CC       by strains causing extraintestinal infections, such as those of
CC       the urinary tract.
CC   -!- SIMILARITY: BELONGS TO THE RTX PROKARYOTIC TOXIN FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; M14107; AAA98233.1; -.
DR   InterPro; IPR001343; Hemlysn_Ca_bind.
DR   InterPro; IPR003995; RtxA.
DR   Pfam; PF00353; hemolysinCabind; 6.
DR   Pfam; PF02382; RTX; 1.
DR   PRINTS; PR00313; CABNDNGRPT.
DR   PRINTS; PR01488; RTXTOXINA.
DR   PROSITE; PS00330; HEMOLYSIN_CALCIUM; 4.
KW   Hemolysis; Toxin; Cytolysis; Repeat; Calcium; Transmembrane;
KW   Lipoprotein; Palmitate; Plasmid.
FT   TRANSMEM    238    260       POTENTIAL.
FT   TRANSMEM    268    327       POTENTIAL.
FT   TRANSMEM    365    411       POTENTIAL.
FT   DOMAIN      724    870       16 X REPEATS, GLY-RICH.
FT   REPEAT      724    729       1.
FT   REPEAT      733    738       2.
FT   REPEAT      742    747       3.
FT   REPEAT      751    756       4.
FT   REPEAT      760    765       5.
FT   REPEAT      769    774       6.
FT   REPEAT      778    783       7.
FT   REPEAT      787    792       8.
FT   REPEAT      796    801       9.
FT   REPEAT      807    812       10.
FT   REPEAT      817    822       11.
FT   REPEAT      826    831       12.
FT   REPEAT      835    840       13.
FT   REPEAT      844    849       14.
FT   REPEAT      856    861       15.
FT   REPEAT      865    870       16.
FT   LIPID       564    564       PALMITATE.
FT   LIPID       690    690       PALMITATE.
SQ   SEQUENCE   1024 AA;  110201 MW;  83944917F76C945B CRC64;

  Query Match             1.5%;  Score 14;  DB 1;  Length 1024;
  Best Local Similarity   100.0%;  Pred. No. 2.2e-05;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        379 ISGILEASKQAMFE 392
              ||||||||||||||
Db        406 ISGILEASKQAMFE 419

RESULT 3
RT31_ACTPL
ID   RT31_ACTPL     STANDARD;      PRT;  1049 AA.
AC   P55130;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   RTX-III toxin determinant A from serotype 2 (APX-IIIA) (Cytolysin
DE   IIIA) (CLY-IIIA).
GN   APXIIIA OR CLYIIIA OR RTXA OR PTXA.
OS   Actinobacillus pleuropneumoniae (Haemophilus pleuropneumoniae).
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC   Pasteurellaceae; Actinobacillus.
OX   NCBI_TaxID=715;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Serotype 2;
RX   MEDLINE=93263992; PubMed=8494611;
RA   Chang Y.-F., Shi J., Ma D.-P., Shin S.J., Lein D.H.;
RT   "Molecular analysis of the Actinobacillus pleuropneumoniae RTX
RT   toxin-III gene cluster.";
RL   DNA Cell Biol. 12:351-362(1993).
RN   [2]
RP   SEQUENCE OF 828-1049 FROM N.A.
RC   STRAIN=1536 / Serotype 2;
RX   MEDLINE=95012630; PubMed=7927703;
RA   Jansen R., Briaire J., van Geel A.B.M., Kamp E.M., Gielkens A.L.J.,
RA   Smits M.A.;
RT   "Genetic map of the Actinobacillus pleuropneumoniae RTX-toxin (Apx)
RT   operons: characterization of the ApxIII operons.";
RL   Infect. Immun. 62:4411-4418(1994).
CC   -!- FUNCTION: Does not have hemolytic activity but shows a strong
CC       cytotoxicity towards alveolar macrophages and neutrophils.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- DOMAIN: The Gly-rich region is probably involved in binding
CC       calcium, which is required for target cell-binding or cytolytic
CC       activity (By similarity).
CC   -!- DOMAIN: The three transmembrane domains are believed to be
CC       involved in pore formation by the cytotoxin (By similarity).
CC   -!- PTM: Palmitoylated by apxIIIC. The toxin only becomes active when
CC       modified (By similarity).
CC   -!- SIMILARITY: BELONGS TO THE RTX PROKARYOTIC TOXIN FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; L12145; AAA21924.1; -.
DR   EMBL; X80056; CAB37652.1; ALT_SEQ.
DR   PIR; S51784; S51784.
DR   InterPro; IPR001343; Hemlysn_Ca_bind.
DR   InterPro; IPR003995; RtxA.
DR   Pfam; PF00353; hemolysinCabind; 6.
DR   Pfam; PF02382; RTX; 1.
DR   PRINTS; PR00313; CABNDNGRPT.
DR   PRINTS; PR01488; RTXTOXINA.
DR   PROSITE; PS00330; HEMOLYSIN_CALCIUM; 3.
KW   Toxin; Cytolysis; Repeat; Calcium; Transmembrane; Lipoprotein;
KW   Palmitate.
FT   TRANSMEM    154    170       POTENTIAL.
FT   TRANSMEM    315    331       POTENTIAL.
FT   TRANSMEM    397    413       POTENTIAL.
FT   DOMAIN      753    858       11 X REPEATS, GLY-RICH.
FT   REPEAT      753    758       1.
FT   REPEAT      762    767       2.
```

```
FT   REPEAT       771    776       3.
FT   REPEAT       780    785       4.
FT   REPEAT       789    794       5.
FT   REPEAT       798    803       6.
FT   REPEAT       807    812       7.
FT   REPEAT       826    831       8.
FT   REPEAT       835    840       9.
FT   REPEAT       844    849       10.
FT   REPEAT       853    858       11.
SQ   SEQUENCE   1049 AA;  112491 MW;  F99846BFD4E5CE72 CRC64;

  Query Match             1.5%;  Score 14;  DB 1;  Length 1049;
  Best Local Similarity   100.0%;  Pred. No. 2.2e-05;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        241 KVAAGFELSNQVIG 254
              ||||||||||||||
Db        275 KVAAGFELSNQVIG 288


RESULT 4
RT32_ACTPL
ID   RT32_ACTPL     STANDARD;      PRT;  1052 AA.
AC   P55131;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   RTX-III toxin determinant A from serotype 8 (APX-IIIA) (Cytolysin
DE   IIIA) (CLY-IIIA).
GN   APXIIIA OR CLYIIIA OR RTXA OR PTXA.
OS   Actinobacillus pleuropneumoniae (Haemophilus pleuropneumoniae).
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC   Pasteurellaceae; Actinobacillus.
OX   NCBI_TaxID=715;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=405 / Serotype 8;
RX   MEDLINE=95012630; PubMed=7927703;
RA   Jansen R., Briaire J., van Geel A.B.M., Kamp E.M., Gielkens A.L.J.,
RA   Smits M.A.;
RT   "Genetic map of the Actinobacillus pleuropneumoniae RTX-toxin (Apx)
RT   operons: characterization of the ApxIII operons.";
RL   Infect. Immun. 62:4411-4418(1994).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Serotype 8;
RX   MEDLINE=93162836; PubMed=8432615;
RA   Jansen R., Briaire J., Kamp E.M., Gielkens A.L.J., Smits M.A.;
RT   "Cloning and characterization of the Actinobacillus
RT   pleuropneumoniae-RTX-toxin III (ApxIII) gene.";
RL   Infect. Immun. 61:947-954(1993).
CC   -!- FUNCTION: Does not have hemolytic activity but shows a strong
CC       cytotoxicity towards alveolar macrophages and neutrophils.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- DOMAIN: The Gly-rich region is probably involved in binding
CC       calcium, which is required for target cell-binding or cytolytic
CC       activity (By similarity).
CC   -!- DOMAIN: The three transmembrane domains are believed to be
CC       involved in pore formation by the cytotoxin (By similarity).
CC   -!- PTM: Palmitoylated by apxIIIC. The toxin only becomes active when
CC       modified (By similarity).
CC   -!- SIMILARITY: BELONGS TO THE RTX PROKARYOTIC TOXIN FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; X80055; CAA56358.1; -.
DR   EMBL; X68815; CAA48711.1; -.
DR   PIR; B49219; B49219.
DR   InterPro; IPR001343; Hemlysn_Ca_bind.
DR   InterPro; IPR003995; RtxA.
DR   Pfam; PF00353; hemolysinCabind; 6.
DR   Pfam; PF02382; RTX; 1.
DR   PRINTS; PR00313; CABNDNGRPT.
DR   PRINTS; PR01488; RTXTOXINA.
DR   PROSITE; PS00330; HEMOLYSIN_CALCIUM; 3.
KW   Toxin; Cytolysis; Repeat; Calcium; Transmembrane; Lipoprotein;
KW   Palmitate.
FT   TRANSMEM     248    265       POTENTIAL.
FT   TRANSMEM     275    334       POTENTIAL.
FT   TRANSMEM     372    418       POTENTIAL.
FT   DOMAIN       754    859       7 X REPEATS, GLY-RICH.
FT   REPEAT       754    759       1.
FT   REPEAT       763    768       2.
FT   REPEAT       772    777       3.
FT   REPEAT       781    786       4.
FT   REPEAT       790    795       5.
FT   REPEAT       799    804       6.
FT   REPEAT       808    813       7.
FT   REPEAT       827    832       8.
FT   REPEAT       836    841       9.
FT   REPEAT       845    850       10.
FT   REPEAT       854    859       11.
SQ   SEQUENCE   1052 AA;  112809 MW;  F83AFE25A6FD8758 CRC64;

  Query Match             1.5%;  Score 14;  DB 1;  Length 1052;
  Best Local Similarity   100.0%;  Pred. No. 2.2e-05;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        241 KVAAGFELSNQVIG 254
              ||||||||||||||
Db        275 KVAAGFELSNQVIG 288


RESULT 5
HLYA_ACTSU
ID   HLYA_ACTSU     STANDARD;      PRT;   956 AA.
AC   Q00951;
DT   01-APR-1993 (Rel. 25, Created)
DT   01-APR-1993 (Rel. 25, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Hemolysin (Cytolysin II) (CLY-IIA) (HLY-IIA) (CYTC) (APPA).
GN   APPA OR CLYIIA OR HLYIIA OR CYTC.
OS   Actinobacillus suis.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC   Pasteurellaceae; Actinobacillus.
OX   NCBI_TaxID=716;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=3714;
RX   MEDLINE=92267623; PubMed=1587585;
RA   Burrows L.L., Lo R.Y.;
RT   "Molecular characterization of an RTX toxin determinant from
RT   Actinobacillus suis.";
RL   Infect. Immun. 60:2166-2173(1992).
CC   -!- FUNCTION: One of the virulence factors of A.suis might be a
CC       secreted cytotoxin, possibly the extracellular hemolysin.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- DOMAIN: The Gly-rich region is probably involved in binding
CC       calcium, which is required for target cell-binding or cytolytic
CC       activity.
CC   -!- DOMAIN: The three transmembrane domains are believed to be
CC       involved in pore formation by the cytotoxin (By similarity).
CC   -!- PTM: Palmitoylated by lktC. The toxin only becomes active when
CC       modified (By similarity).
CC   -!- SIMILARITY: BELONGS TO THE RTX PROKARYOTIC TOXIN FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
```

```
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----------------------------------------------------------------------
DR   EMBL; M90440; AAA21918.1; -.
DR   InterPro; IPR001343; Hemlysn_Ca_bind.
DR   InterPro; IPR003995; RtxA.
DR   Pfam; PF00353; hemolysinCabind; 5.
DR   Pfam; PF02382; RTX; 1.
DR   PRINTS; PR00313; CABNDNGRPT.
DR   PRINTS; PR01488; RTXTOXINA.
DR   PROSITE; PS00330; HEMOLYSIN_CALCIUM; 1.
KW   Hemolysis; Toxin; Cytolysis; Repeat; Calcium; Transmembrane;
KW   Lipoprotein; Palmitate.
FT   TRANSMEM    238    254       POTENTIAL.
FT   TRANSMEM    302    320       POTENTIAL.
FT   TRANSMEM    383    406       POTENTIAL.
FT   DOMAIN      719    799       9 X REPEATS, GLY-RICH.
FT   REPEAT      719    724       1.
FT   REPEAT      728    733       2.
FT   REPEAT      737    742       3.
FT   REPEAT      746    751       4.
FT   REPEAT      755    760       5.
FT   REPEAT      764    769       6.
FT   REPEAT      773    778       7.
FT   REPEAT      782    787       8.
FT   REPEAT      794    799       9.
SQ   SEQUENCE   956 AA;  102453 MW;  3415FF1D7ADD4365 CRC64;

  Query Match             1.2%;  Score 11;  DB 1;  Length 956;
  Best Local Similarity   100.0%;  Pred. No. 0.023;
  Matches   11;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        369 ALLVAGVTGLI 379
              |||||||||||
Db        392 ALLVAGVTGLI 402


RESULT 6
RT2A_ACTPL
ID   RT2A_ACTPL     STANDARD;      PRT;   956 AA.
AC   P15377;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-APR-1990 (Rel. 14, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   RTX-II toxin determinant A (APX-IIA) (Hemolysin IIA) (HLY-IIA)
DE   (Cytolysin IIA) (CLY-IIA).
GN   APXIIA OR CLYIIA OR HLYIIA OR APPA OR CYTC.
OS   Actinobacillus pleuropneumoniae (Haemophilus pleuropneumoniae).
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC   Pasteurellaceae; Actinobacillus.
OX   NCBI_TaxID=715;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Serotype 5;
RX   MEDLINE=90126233; PubMed=2693022;
RA   Chang Y.-F., Young R., Struck D.K.;
RT   "Cloning and characterization of a hemolysin gene from Actinobacillus
RT   (Haemophilus) pleuropneumoniae.";
RL   DNA 8:635-647(1989).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Isolate CVI 13261 / Serotype 9;
RX   MEDLINE=92040145; PubMed=1937809;
RA   Smits M.A., Briaire J., Jansen R., Smith H.E., Kamp E.M.,
RA   Gielkens A.L.;
RT   "Cytolysins of Actinobacillus pleuropneumoniae serotype 9.";
RL   Infect. Immun. 59:4497-4504(1991).
CC   -!- FUNCTION: One of the virulence factors of A.pleuropneumoniae,
CC       which shows a weak hemolytic activity and is moderately cytotoxic
CC       for alveolar macrophages and neutrophils.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- DOMAIN: The Gly-rich region is probably involved in binding
CC       calcium, which is required for target cell-binding or cytolytic
CC       activity (By similarity).
CC   -!- DOMAIN: The three transmembrane domains are believed to be
CC       involved in pore formation by the cytotoxin (By similarity).
CC   -!- PTM: Palmitoylated by apxIIC. The toxin only becomes active when
CC       modified (By similarity).
CC   -!- SIMILARITY: BELONGS TO THE RTX PROKARYOTIC TOXIN FAMILY.
CC   -----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----------------------------------------------------------------------
DR   EMBL; M30602; AAA87232.1; -.
DR   EMBL; X61111; CAA43423.1; -.
DR   PIR; B33389; B33389.
DR   InterPro; IPR001343; Hemlysn_Ca_bind.
DR   InterPro; IPR003995; RtxA.
DR   Pfam; PF00353; hemolysinCabind; 5.
DR   Pfam; PF02382; RTX; 1.
DR   PRINTS; PR00313; CABNDNGRPT.
DR   PRINTS; PR01488; RTXTOXINA.
DR   PROSITE; PS00330; HEMOLYSIN_CALCIUM; 1.
KW   Hemolysis; Toxin; Cytolysis; Repeat; Calcium; Transmembrane;
KW   Lipoprotein; Palmitate.
FT   TRANSMEM    233    256       POTENTIAL.
FT   TRANSMEM    266    323       POTENTIAL.
FT   TRANSMEM    361    406       POTENTIAL.
FT   DOMAIN      719    787       8 X REPEATS, GLY-RICH.
FT   REPEAT      719    724       1.
FT   REPEAT      728    733       2.
FT   REPEAT      737    742       3.
FT   REPEAT      746    751       4.
FT   REPEAT      755    760       5.
FT   REPEAT      764    769       6.
FT   REPEAT      773    778       7.
FT   REPEAT      782    787       8.
SQ   SEQUENCE   956 AA;  102531 MW;  BDBCABBADF14A641 CRC64;

  Query Match             1.2%;  Score 11;  DB 1;  Length 956;
  Best Local Similarity   100.0%;  Pred. No. 0.023;
  Matches   11;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        369 ALLVAGVTGLI 379
              |||||||||||
Db        392 ALLVAGVTGLI 402


RESULT 7
RT11_ACTPL
ID   RT11_ACTPL     STANDARD;      PRT;  1023 AA.
AC   P55128;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   RTX-I toxin determinant A from serotypes 1/9 (APX-IA) (Hemolysin IA)
DE   (HLY-IA) (Cytolysin IA) (CLY-IA).
GN   APXIA OR CLYIA OR HLYIA.
OS   Actinobacillus pleuropneumoniae (Haemophilus pleuropneumoniae).
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC   Pasteurellaceae; Actinobacillus.
OX   NCBI_TaxID=715;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=S 4074 / Serotype 1;
RX   MEDLINE=91348845; PubMed=1879928;
RA   Frey J., Meier R., Gygi D., Nicolet J.;
RT   "Nucleotide sequence of the hemolysin I gene from Actinobacillus
RT   pleuropneumoniae.";
RL   Infect. Immun. 59:3026-3032(1991).
```

```
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=S 4074 / Serotype 1;
RX   MEDLINE=94237497; PubMed=8181764;
RA   Frey J., Haldimann A., Nicolet J., Boffini A., Prentki P.;
RT   "Sequence analysis and transcription of the apxI operon (hemolysin I)
RT   from Actinobacillus pleuropneumoniae.";
RL   Gene 142:97-102(1994).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Isolate CVI 13261 / Serotype 9;
RX   MEDLINE=93366425; PubMed=8359891;
RA   Jansen R., Briaire J., Kamp E.M., Gielkens A.L.J., Smits M.A.;
RT   "Structural analysis of the Actinobacillus pleuropneumoniae-RTX-toxin
RT   I (ApxI) operon.";
RL   Infect. Immun. 61:3688-3695(1993).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=S 4074 / Serotype 1;
RA   Chang Y., Wang Y., Chin N.;
RL   Submitted (JAN-1994) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: One of the virulence factors of A.pleuropneumoniae,
CC       which has a strong hemolytic activity and is cytotoxic for
CC       alveolar macrophages and neutrophils.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- DOMAIN: The Gly-rich region is probably involved in binding
CC       calcium, which is required for target cell-binding or cytolytic
CC       activity.
CC   -!- DOMAIN: The three transmembrane domains are believed to be
CC       involved in pore formation by the cytotoxin (By similarity).
CC   -!- PTM: Palmitoylated by apxIC. The toxin only becomes active when
CC       modified (By similarity).
CC   -!- MISCELLANEOUS: ApxIA is partially deleted in serotypes 2, 4, 6, 7,
CC       8, 12, and totally deleted in serotype 3.
CC   -!- MISCELLANEOUS: The sequence shown is that of serotype 1.
CC   -!- SIMILARITY: BELONGS TO THE RTX PROKARYOTIC TOXIN FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; X52899; CAA37081.1; -.
DR   EMBL; X68595; CAA48586.1; -.
DR   EMBL; X73117; CAA51548.1; -.
DR   EMBL; U05042; AAB05034.1; -.
DR   PIR; I39643; I39643.
DR   InterPro; IPR001343; Hemlysn_Ca_bind.
DR   InterPro; IPR003995; RtxA.
DR   Pfam; PF00353; hemolysinCabind; 6.
DR   Pfam; PF02382; RTX; 1.
DR   PRINTS; PR00313; CABNDNGRPT.
DR   PRINTS; PR01488; RTXTOXINA.
DR   PROSITE; PS00330; HEMOLYSIN_CALCIUM; 2.
KW   Hemolysis; Toxin; Cytolysis; Repeat; Calcium; Transmembrane;
KW   Lipoprotein; Palmitate.
FT   TRANSMEM     226    256       POTENTIAL.
FT   TRANSMEM     297    326       POTENTIAL.
FT   TRANSMEM     367    406       POTENTIAL.
FT   DOMAIN       722    845       13 X REPEATS, GLY-RICH.
FT   REPEAT       722    727       1.
FT   REPEAT       731    736       2.
FT   REPEAT       740    745       3.
FT   REPEAT       749    754       4.
FT   REPEAT       758    763       5.
FT   REPEAT       767    772       6.
FT   REPEAT       776    781       7.
FT   REPEAT       785    790       8.
FT   REPEAT       794    799       9.
FT   REPEAT       813    818       10.
FT   REPEAT       822    827       11.
FT   REPEAT       831    836       12.
FT   REPEAT       840    845       13.
FT   CONFLICT     210    217       AMPYLTLA -> GNALSNTR (IN REF. 3 AND 4).
FT   CONFLICT     374    374       R -> A (IN REF. 3 AND 4).
FT   CONFLICT     562    562       Q -> E (IN REF. 3 AND 4).
FT   CONFLICT     687    688       TC -> R (IN REF. 3 AND 4).
SQ   SEQUENCE   1023 AA;  110193 MW;  F99A88CFC9F1A598 CRC64;

  Query Match             1.2%;  Score 11;  DB 1;  Length 1023;
  Best Local Similarity   100.0%;  Pred. No. 0.024;
  Matches   11;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        266 AQRVAAGLSTT 276
              |||||||||||
Db        289 AQRVAAGLSTT 299


RESULT 8
RT12_ACTPL
ID   RT12_ACTPL     STANDARD;      PRT;  1023 AA.
AC   P55129;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   RTX-I toxin determinant A from serotypes 5/10 (APX-IA) (Hemolysin IA)
DE   (HLY-IA) (Cytolysin IA) (CLY-IA).
GN   APXIA OR CLYIA OR HLYIA.
OS   Actinobacillus pleuropneumoniae (Haemophilus pleuropneumoniae).
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC   Pasteurellaceae; Actinobacillus.
OX   NCBI_TaxID=715;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=13039 / Serotype 10;
RX   MEDLINE=94276858; PubMed=8007819;
RA   Nagai S., Yagihashi T., Ishihama A.;
RT   "DNA sequence analysis of an allelic variant of the Actinobacillus
RT   pleuropneumoniae-RTX-toxin I (ApxIA) from serotype 10.";
RL   Microb. Pathog. 15:485-495(1993).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K17 / Serotype 5;
RX   MEDLINE=96401417; PubMed=8807793;
RA   Chin N., Frey J., Chang C.F., Chang Y.F.;
RT   "Identification of a locus involved in the utilization of iron by
RT   Actinobacillus pleuropneumoniae.";
RL   FEMS Microbiol. Lett. 143:1-6(1996).
RN   [3]
RP   SEQUENCE OF 886-1023 FROM N.A.
RC   STRAIN=K17 / Serotype 5;
RX   MEDLINE=93366425; PubMed=8359891;
RA   Jansen R., Briaire J., Kamp E.M., Gielkens A.L.J., Smits M.A.;
RT   "Structural analysis of the Actinobacillus pleuropneumoniae-RTX-toxin
RT   I (ApxI) operon.";
RL   Infect. Immun. 61:3688-3695(1993).
CC   -!- FUNCTION: One of the virulence factors of A.pleuropneumoniae,
CC       which has a strong hemolytic activity and is cytotoxic for
CC       alveolar macrophages and neutrophils.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- DOMAIN: The Gly-rich region is probably involved in binding
CC       calcium, which is required for target cell-binding or cytolytic
CC       activity.
CC   -!- DOMAIN: The three transmembrane domains are believed to be
CC       involved in pore formation by the cytotoxin (By similarity).
CC   -!- PTM: PALMITOYLATED BY APXIC. The toxin only becomes active when
CC       modified (By similarity).
CC   -!- MISCELLANEOUS: ApxIA is partially deleted in serotypes 2, 4, 6, 7,
CC       8, 12, and totally deleted in serotype 3.
CC   -!- MISCELLANEOUS: The sequence shown is that of serotype 10.
CC   -!- SIMILARITY: BELONGS TO THE RTX PROKARYOTIC TOXIN FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
```

```
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----------------------------------------------------------------------
DR   EMBL; D16582; BAA04014.1; -.
DR   EMBL; U04954; AAB17220.1; -.
DR   EMBL; X73116; CAA51546.1; -.
DR   PIR; I39641; I39641.
DR   InterPro; IPR001343; Hemlysn_Ca_bind.
DR   InterPro; IPR003995; RtxA.
DR   Pfam; PF00353; hemolysinCabind; 6.
DR   Pfam; PF02382; RTX; 1.
DR   PRINTS; PR00313; CABNDNGRPT.
DR   PRINTS; PR01488; RTXTOXINA.
DR   PROSITE; PS00330; HEMOLYSIN_CALCIUM; 2.
KW   Hemolysis; Toxin; Cytolysis; Repeat; Calcium; Transmembrane;
KW   Lipoprotein; Palmitate.
FT   TRANSMEM    226    256       POTENTIAL.
FT   TRANSMEM    297    326       POTENTIAL.
FT   TRANSMEM    367    406       POTENTIAL.
FT   DOMAIN      722    845       13 X REPEATS, GLY-RICH.
FT   REPEAT      722    727       1.
FT   REPEAT      731    736       2.
FT   REPEAT      740    745       3.
FT   REPEAT      749    754       4.
FT   REPEAT      758    763       5.
FT   REPEAT      767    772       6.
FT   REPEAT      776    781       7.
FT   REPEAT      785    790       8.
FT   REPEAT      794    799       9.
FT   REPEAT      813    818       10.
FT   REPEAT      822    827       11.
FT   REPEAT      831    836       12.
FT   REPEAT      840    845       13.
FT   CONFLICT    210    217       AMPYLTLA -> GNALSNTR (IN REF. 2).
FT   CONFLICT    581    581       E -> Q (IN REF. 2).
FT   CONFLICT    687    688       TC -> R (IN REF. 2).
FT   CONFLICT   1015   1015       F -> L (IN REF. 2).
SQ   SEQUENCE   1023 AA;  110129 MW;  183C7C15EE57DB55 CRC64;

  Query Match             1.2%;  Score 11;  DB 1;  Length 1023;
  Best Local Similarity   100.0%;  Pred. No. 0.024;
  Matches   11;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        266 AQRVAAGLSTT 276
              |||||||||||
Db        289 AQRVAAGLSTT 299

RESULT 9
LKTA_PASSP
ID   LKTA_PASSP     STANDARD;      PRT;   947 AA.
AC   P55123;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Leukotoxin.
GN   LKTA.
OS   Pasteurella haemolytica-like sp. (strain 5943B).
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC   Pasteurellaceae; Pasteurella.
OX   NCBI_TaxID=28165;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93239320; PubMed=8478098;
RA   Chang Y.-F., Ma D.-P., Shi J., Chengappa M.M.;
RT   "Molecular characterization of a leukotoxin gene from a Pasteurella
RT   haemolytica-like organism, encoding a new member of the RTX toxin
RT   family.";
RL   Infect. Immun. 61:2089-2095(1993).
```

```
CC   -!- FUNCTION: Virulence factor which is cytotoxic for leukocytes but
CC       is not hemolytic.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- DOMAIN: The Gly-rich region is probably involved in binding
CC       calcium, which is required for target cell-binding or cytolytic
CC       activity.
CC   -!- DOMAIN: The three transmembrane domains are believed to be
CC       involved in pore formation by the cytotoxin (By similarity).
CC   -!- PTM: Palmitoylated by lktC. The toxin only becomes active when
CC       modified (By similarity).
CC   -!- SIMILARITY: BELONGS TO THE RTX PROKARYOTIC TOXIN FAMILY.
CC   -----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----------------------------------------------------------------------
DR   EMBL; L12148; AAA16444.1; -.
DR   InterPro; IPR001343; Hemlysn_Ca_bind.
DR   InterPro; IPR003995; RtxA.
DR   Pfam; PF00353; hemolysinCabind; 5.
DR   Pfam; PF02382; RTX; 1.
DR   PRINTS; PR00313; CABNDNGRPT.
DR   PRINTS; PR01488; RTXTOXINA.
DR   PROSITE; PS00330; HEMOLYSIN_CALCIUM; 4.
KW   Hemolysis; Toxin; Cytolysis; Repeat; Calcium; Transmembrane;
KW   Lipoprotein; Palmitate.
FT   TRANSMEM    154    170       POTENTIAL.
FT   TRANSMEM    312    333       POTENTIAL.
FT   TRANSMEM    393    414       POTENTIAL.
FT   DOMAIN      625    780       7 X REPEATS, GLY-RICH.
FT   REPEAT      625    630       1.
FT   REPEAT      730    735       2.
FT   REPEAT      739    744       3.
FT   REPEAT      748    753       4.
FT   REPEAT      757    762       5.
FT   REPEAT      766    771       6.
FT   REPEAT      775    780       7.
SQ   SEQUENCE   947 AA;  101559 MW;  9744F06395EF5BED CRC64;

  Query Match             1.1%;  Score 10;  DB 1;  Length 947;
  Best Local Similarity   100.0%;  Pred. No. 0.23;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        265 LAQRVAAGLS 274
              ||||||||||
Db        279 LAQRVAAGLS 288

RESULT 10
LKA1_PASHA
ID   LKA1_PASHA     STANDARD;      PRT;   953 AA.
AC   P16535;
DT   01-AUG-1990 (Rel. 15, Created)
DT   01-AUG-1990 (Rel. 15, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Leukotoxin from serotype A1.
GN   LKTA.
OS   Pasteurella haemolytica.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC   Pasteurellaceae; Mannheimia.
OX   NCBI_TaxID=75985;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Serotype A1;
RX   MEDLINE=87306837; PubMed=3040588;
RA   Lo R.Y.C., Strathdee C.A., Shewen P.E.;
RT   "Nucleotide sequence of the leukotoxin genes of Pasteurella
RT   haemolytica A1.";
RL   Infect. Immun. 55:1987-1996(1987).
```

```
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Serotype A1 / PHL101;
RX   MEDLINE=89210283; PubMed=2707120;
RA   Highlander S.K., Chidambaram M., Engler M.J., Weinstock G.M.;
RT   "DNA sequence of the Pasteurella haemolytica leukotoxin gene
RT   cluster.";
RL   DNA 8:15-28(1989).
RN   [3]
RP   SEQUENCE OF 884-953 FROM N.A.
RC   STRAIN=Serotype A1 / PHL101;
RX   MEDLINE=90236888; PubMed=2185213;
RA   Highlander S.K., Engler M.J., Weinstock G.M.;
RT   "Secretion and expression of the Pasteurella haemolytica Leukotoxin.";
RL   J. Bacteriol. 172:2343-2350(1990).
CC   -!- FUNCTION: Bacterial hemolysins are exotoxins that attack blood
CC       cell membranes and cause cell rupture by mechanisms not clearly
CC       defined.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- DOMAIN: The Gly-rich region is probably involved in binding
CC       calcium, which is required for target cell-binding or cytolytic
CC       activity.
CC   -!- DOMAIN: The three transmembrane domains are believed to be
CC       involved in pore formation by the cytotoxin (By similarity).
CC   -!- PTM: Palmitoylated by lktC. The toxin only becomes active when
CC       modified (By similarity).
CC   -!- SIMILARITY: BELONGS TO THE RTX PROKARYOTIC TOXIN FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; M20730; AAA25529.1; -.
DR   EMBL; M24197; AAA25543.1; -.
DR   PIR; B30169; B30169.
DR   InterPro; IPR001343; Hemlysn_Ca_bind.
DR   InterPro; IPR003995; RtxA.
DR   Pfam; PF00353; hemolysinCabind; 5.
DR   Pfam; PF02382; RTX; 1.
DR   PRINTS; PR00313; CABNDNGRPT.
DR   PRINTS; PR01488; RTXTOXINA.
DR   PROSITE; PS00330; HEMOLYSIN_CALCIUM; 4.
KW   Hemolysis; Toxin; Cytolysis; Repeat; Calcium; Transmembrane;
KW   Lipoprotein; Palmitate.
FT   TRANSMEM    229    249       POTENTIAL.
FT   TRANSMEM    297    318       POTENTIAL.
FT   TRANSMEM    366    390       POTENTIAL.
FT   DOMAIN      734    784       6 X REPEATS, GLY-RICH (BY SIMILARITY).
FT   REPEAT      734    739       1.
FT   REPEAT      743    748       2.
FT   REPEAT      752    757       3.
FT   REPEAT      761    766       4.
FT   REPEAT      770    775       5.
FT   REPEAT      779    784       6.
FT   CONFLICT    409    414       FEHVAN -> LSTLQI (IN REF. 2).
FT   CONFLICT    742    742       D -> Y (IN REF. 2).
SQ   SEQUENCE   953 AA;  101996 MW;  7F93D113A118C05F CRC64;

  Query Match             1.1%;  Score 10;  DB 1;  Length 953;
  Best Local Similarity   100.0%;  Pred. No. 0.23;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        265 LAQRVAAGLS 274
              ||||||||||
Db        283 LAQRVAAGLS 292


RESULT 11
LKA3_PASHA
```

```
ID   LKA3_PASHA     STANDARD;      PRT;   953 AA.
AC   P55116;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Leukotoxin from serotype T3.
GN   LKTA.
OS   Pasteurella haemolytica.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC   Pasteurellaceae; Mannheimia.
OX   NCBI_TaxID=75985;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Serotype T3;
RX   MEDLINE=94041617; PubMed=8225575;
RA   Burrows L.L., Lo R.Y.C., Olah-Winfield E.;
RT   "Molecular analysis of the leukotoxin determinants from Pasteurella
RT   haemolytica serotypes 1 to 16.";
RL   Infect. Immun. 61:5001-5007(1993).
CC   -!- FUNCTION: Bacterial hemolysins are exotoxins that attack blood
CC       cell membranes and cause cell rupture by mechanisms not clearly
CC       defined.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- DOMAIN: The Gly-rich region is probably involved in binding
CC       calcium, which is required for target cell-binding or cytolytic
CC       activity.
CC   -!- DOMAIN: The three transmembrane domains are believed to be
CC       involved in pore formation by the cytotoxin (BY SIMILARITY).
CC   -!- PTM: Palmitoylated by lktC. The toxin only becomes active when
CC       modified (By similarity).
CC   -!- SIMILARITY: BELONGS TO THE RTX PROKARYOTIC TOXIN FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; U01216; AAB36691.1; -.
DR   InterPro; IPR001343; Hemlysn_Ca_bind.
DR   InterPro; IPR003995; RtxA.
DR   Pfam; PF00353; hemolysinCabind; 5.
DR   Pfam; PF02382; RTX; 1.
DR   PRINTS; PR00313; CABNDNGRPT.
DR   PRINTS; PR01488; RTXTOXINA.
DR   PROSITE; PS00330; HEMOLYSIN_CALCIUM; 2.
KW   Hemolysis; Toxin; Cytolysis; Repeat; Calcium; Transmembrane;
KW   Lipoprotein; Palmitate.
FT   TRANSMEM    229    249       POTENTIAL.
FT   TRANSMEM    297    318       POTENTIAL.
FT   TRANSMEM    381    401       POTENTIAL.
FT   DOMAIN      734    784       6 X REPEATS, GLY-RICH.
FT   REPEAT      734    739       1.
FT   REPEAT      743    748       2.
FT   REPEAT      752    757       3.
FT   REPEAT      761    766       4.
FT   REPEAT      770    775       5.
FT   REPEAT      779    784       6.
SQ   SEQUENCE   953 AA;  101948 MW;  FDBDCE2FDC85FDF2 CRC64;

  Query Match             1.1%;  Score 10;  DB 1;  Length 953;
  Best Local Similarity   100.0%;  Pred. No. 0.23;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        443 AERVIAITQQ 452
              ||||||||||
Db        461 AERVIAITQQ 470


RESULT 12
LKAB_PASHA
```

```
ID   LKAB_PASHA     STANDARD;      PRT;   953 AA.
AC   P55118;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Leukotoxin from serotype A11.
GN   LKTA.
OS   Pasteurella haemolytica.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC   Pasteurellaceae; Mannheimia.
OX   NCBI_TaxID=75985;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Serotype A11;
RX   MEDLINE=94041617; PubMed=8225575;
RA   Burrows L.L., Olah-Winfield E., Lo R.Y.C.;
RT   "Molecular analysis of the leukotoxin determinants from Pasteurella
RT   haemolytica serotypes 1 to 16.";
RL   Infect. Immun. 61:5001-5007(1993).
CC   -!- FUNCTION: Bacterial hemolysins are exotoxins that attack blood
CC       cell membranes and cause cell rupture by mechanisms not clearly
CC       DEFINED.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- DOMAIN: The Gly-rich region is probably involved in binding
CC       calcium, which is required for target cell-binding or cytolytic
CC       activity.
CC   -!- DOMAIN: The three transmembrane domains are believed to be
CC       involved in pore formation by the cytotoxin (BY SIMILARITY).
CC   -!- PTM: Palmitoylated by lktC. The toxin only becomes active when
CC       modified (By similarity).
CC   -!- SIMILARITY: BELONGS TO THE RTX PROKARYOTIC TOXIN FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; U01215; AAB36689.1; -.
DR   InterPro; IPR001343; Hemlysn_Ca_bind.
DR   InterPro; IPR003995; RtxA.
DR   Pfam; PF00353; hemolysinCabind; 5.
DR   Pfam; PF02382; RTX; 1.
DR   PRINTS; PR00313; CABNDNGRPT.
DR   PRINTS; PR01488; RTXTOXINA.
DR   PROSITE; PS00330; HEMOLYSIN_CALCIUM; 4.
KW   Hemolysis; Toxin; Cytolysis; Repeat; Calcium; Transmembrane;
KW   Lipoprotein; Palmitate.
FT   TRANSMEM    230    250       POTENTIAL.
FT   TRANSMEM    297    317       POTENTIAL.
FT   TRANSMEM    381    401       POTENTIAL.
FT   DOMAIN      734    784       6 X REPEATS, GLY-RICH.
FT   REPEAT      734    739       1.
FT   REPEAT      743    748       2.
FT   REPEAT      752    757       3.
FT   REPEAT      761    766       4.
FT   REPEAT      770    775       5.
FT   REPEAT      779    784       6.
SQ   SEQUENCE   953 AA;  102206 MW;  927FF56CFC884F12 CRC64;

  Query Match             1.1%;  Score 10;  DB 1;  Length 953;
  Best Local Similarity   100.0%;  Pred. No. 0.23;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        265 LAQRVAAGLS 274
              ||||||||||
Db        283 LAQRVAAGLS 292

RESULT 13
LKAA_PASHA
ID   LKAA_PASHA     STANDARD;      PRT;   955 AA.
AC   P55117;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Leukotoxin from serotype T10.
GN   LKTA.
OS   Pasteurella haemolytica.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC   Pasteurellaceae; Mannheimia.
OX   NCBI_TaxID=75985;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Serotype T10;
RX   MEDLINE=96425875; PubMed=8828217;
RA   Lainson F.A., Murray J., Davies R.C., Donachie W.;
RT   "Characterization of epitopes involved in the neutralization of
RT   Pasteurella haemolytica serotype A1 leukotoxin.";
RL   Microbiology 142:2499-2507(1996).
CC   -!- FUNCTION: Bacterial hemolysins are exotoxins that attack blood
CC       cell membranes and cause cell rupture by mechanisms not clearly
CC       defined.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- DOMAIN: The Gly-rich region is probably involved in binding
CC       calcium, which is required for target cell-binding or cytolytic
CC       activity.
CC   -!- DOMAIN: The three transmembrane domains are believed to be
CC       involved in pore formation by the cytotoxin (BY SIMILARITY).
CC   -!- PTM: Palmitoylated by lktC. The toxin only becomes active when
CC       MODIFIED (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE RTX PROKARYOTIC TOXIN FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; Z26247; CAA81206.1; -.
DR   PIR; S37145; A35254.
DR   InterPro; IPR001343; Hemlysn_Ca_bind.
DR   InterPro; IPR003995; RtxA.
DR   Pfam; PF00353; hemolysinCabind; 5.
DR   Pfam; PF02382; RTX; 1.
DR   PRINTS; PR00313; CABNDNGRPT.
DR   PRINTS; PR01488; RTXTOXINA.
DR   PROSITE; PS00330; HEMOLYSIN_CALCIUM; 4.
KW   Hemolysis; Toxin; Cytolysis; Repeat; Calcium; Transmembrane;
KW   Lipoprotein; Palmitate.
FT   TRANSMEM    299    319       POTENTIAL.
FT   TRANSMEM    361    381       POTENTIAL.
FT   TRANSMEM    383    403       POTENTIAL.
FT   DOMAIN      736    786       6 X REPEATS, GLY-RICH.
FT   REPEAT      736    741       1.
FT   REPEAT      745    750       2.
FT   REPEAT      754    759       3.
FT   REPEAT      763    768       4.
FT   REPEAT      772    777       5.
FT   REPEAT      781    786       6.
SQ   SEQUENCE   955 AA;  102187 MW;  B60F2DB8168EBCAF CRC64;

  Query Match             1.1%;  Score 10;  DB 1;  Length 955;
  Best Local Similarity   100.0%;  Pred. No. 0.23;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        265 LAQRVAAGLS 274
              ||||||||||
Db        285 LAQRVAAGLS 294

RESULT 14
```

```
YF14_MYCTU
ID   YF14_MYCTU     STANDARD;      PRT;   262 AA.
AC   P71793;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Putative glycosyl transferase Rv1514c (EC 2.-.-.-).
GN   RV1514C OR MT1564 OR MTCY277.36C.
OS   Mycobacterium tuberculosis.
OC   Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC   Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX   NCBI_TaxID=1773;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=H37Rv;
RX   MEDLINE=98295987; PubMed=9634230;
RA   Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA   Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekaia F.,
RA   Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA   Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
RA   Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA   Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA   Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA   Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RT   "Deciphering the biology of Mycobacterium tuberculosis from the
RT   complete genome sequence.";
RL   Nature 393:537-544(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CDC 1551 / Oshkosh;
RA   Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA   Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
RA   Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
RA   Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA   Bishai W.;
RT   "Whole genome comparison of Mycobacterium tuberculosis clinical and
RT   laboratory strains.";
RL   Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 2.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; Z79701; CAB02028.1; -.
DR   EMBL; AE007023; AAK45831.1; -.
DR   PIR; E70714; E70714.
DR   TIGR; MT1564; -.
DR   TubercuList; Rv1514c; -.
DR   InterPro; IPR001173; Glyco_trans_2.
DR   Pfam; PF00535; Glycos_transf_2; 1.
KW   Hypothetical protein; Transferase; Glycosyltransferase;
KW   Complete proteome.
SQ   SEQUENCE   262 AA;  28965 MW;  6B29BF8D31923E75 CRC64;

  Query Match             0.9%;  Score 8;  DB 1;  Length 262;
  Best Local Similarity   100.0%;  Pred. No. 8.2;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        746 DGGSGDDV 753
              ||||||||
Db         42 DGGSGDDV 49


RESULT 15
LIPA_CAUCR
ID   LIPA_CAUCR     STANDARD;      PRT;   325 AA.
AC   Q9A7I8;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Lipoic acid synthetase (Lip-syn) (Lipoate synthase).
GN   LIPA OR CC1735.
OS   Caulobacter crescentus.
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Caulobacterales;
OC   Caulobacteraceae; Caulobacter.
OX   NCBI_TaxID=155892;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 19089 / CB15;
RX   MEDLINE=21173698; PubMed=11259647;
RA   Nierman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,
RA   Eisen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,
RA   Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,
RA   DeBoy R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H.,
RA   Kolonay J.F., Smit J., Craven M.B., Khouri H., Shetty J., Berry K.,
RA   Utterback T., Tran K., Wolf A., Vamathevan J., Ermolaeva M., White O.,
RA   Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;
RT   "Complete genome sequence of Caulobacter crescentus.";
RL   Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).
CC   -!- FUNCTION: Synthesis of alpha-(+)-lipoic acid. It may be involved
CC       in the sulfur insertion chemistry in lipoate biosynthesis (By
CC       similarity).
CC   -!- PATHWAY: Lipoate biosynthesis.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
CC   -!- SIMILARITY: BELONGS TO THE BIOTIN AND LIPOIC ACID SYNTHETASES
CC       FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; AE005847; AAK23711.1; -.
DR   PIR; C87464; C87464.
DR   TIGR; CC1735; -.
DR   HAMAP; MF_00206; -; 1.
DR   InterPro; IPR006638; Elp3.
DR   InterPro; IPR003698; Lipoate_synth.
DR   Pfam; PF04055; Radical_SAM; 1.
DR   SMART; SM00729; Elp3; 1.
DR   TIGRFAMs; TIGR00510; lipA; 1.
KW   Iron-sulfur; Complete proteome.
FT   METAL        90     90       IRON-SULFUR (POTENTIAL).
FT   METAL        94     94       IRON-SULFUR (POTENTIAL).
FT   METAL        97     97       IRON-SULFUR (POTENTIAL).
SQ   SEQUENCE   325 AA;  36161 MW;  2A1606CD9C3B6400 CRC64;

  Query Match             0.9%;  Score 8;  DB 1;  Length 325;
  Best Local Similarity   100.0%;  Pred. No. 9.9;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        715 DDLLDGGA 722
              ||||||||
Db        135 DDLLDGGA 142


RESULT 16
Y384_MYCGE
ID   Y384_MYCGE     STANDARD;      PRT;   433 AA.
AC   P47624;
DT   01-FEB-1996 (Rel. 33, Created)
DT   01-FEB-1996 (Rel. 33, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Probable GTP-binding protein MG384.
GN   MG384.
OS   Mycoplasma genitalium.
OC   Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX   NCBI_TaxID=2097;
```

```
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 33530 / G-37;
RX   MEDLINE=96026346; PubMed=7569993;
RA   Fraser C.M., Gocayne J.D., White O., Adams M.D., Clayton R.A.,
RA   Fleischmann R.D., Bult C.J., Kerlavage A.R., Sutton G., Kelley J.M.,
RA   Fritchman J.L., Weidman J.F., Small K.V., Sandusky M., Fuhrmann J.L.,
RA   Nguyen D.T., Utterback T.R., Saudek D.M., Phillips C.A., Merrick J.M.,
RA   Tomb J.-F., Dougherty B.A., Bott K.F., Hu P.-C., Lucier T.S.,
RA   Peterson S.N., Smith H.O., Hutchison C.A. III, Venter J.C.;
RT   "The minimal gene complement of Mycoplasma genitalium.";
RL   Science 270:397-403(1995).
CC   -!- SIMILARITY: Belongs to the GTP1 / OBG family.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; U39720; AAC71611.1; -.
DR   PIR; E64242; E64242.
DR   TIGR; MG384; -.
DR   InterPro; IPR006074; GTP1/OBG_dom.
DR   InterPro; IPR006073; GTP1_OBG.
DR   InterPro; IPR006169; GTP1_OBG_sub.
DR   InterPro; IPR005225; Small_GTP.
DR   Pfam; PF01018; GTP1_OBG; 1.
DR   PRINTS; PR00326; GTP1OBG.
DR   TIGRFAMs; TIGR00231; small_GTP; 1.
DR   PROSITE; PS00905; GTP1_OBG; 1.
KW   Hypothetical protein; GTP-binding; Complete proteome.
FT   NP_BIND     166    173       GTP (BY SIMILARITY).
FT   NP_BIND     212    216       GTP (BY SIMILARITY).
FT   NP_BIND     282    285       GTP (BY SIMILARITY).
SQ   SEQUENCE   433 AA;  48166 MW;  08BFC7BC794BC3BE CRC64;

  Query Match              0.9%;  Score 8;  DB 1;  Length 433;
  Best Local Similarity   100.0%;  Pred. No. 13;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         90 KLEKFLQK 97
              ||||||||
Db        295 KLEKFLQK 302


RESULT 17
EGLC_RHIME
ID   EGLC_RHIME     STANDARD;      PRT;   465 AA.
AC   Q9Z3Q2;
DT   30-MAY-2000 (Rel. 39, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Endo-1,3-1,4-beta-glycanase eglC (EC 3.2.1.-) (Succinoglycan
DE   biosynthesis protein eglC).
GN   EGLC OR RA0864 OR SMA1587.
OS   Rhizobium meliloti (Sinorhizobium meliloti).
OG   Plasmid pSymA (megaplasmid 1).
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC   Rhizobiaceae; Sinorhizobium/Ensifer group; Sinorhizobium.
OX   NCBI_TaxID=382;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CXM1-105;
RX   MEDLINE=99413305; PubMed=10485295;
RA   Sharypova L.A., Yurgel S.N., Keller M., Simarov B.V., Puehler A.,
RA   Becker A.;
RT   "The eff-482 locus of Sinorhizobium meliloti CXM1-105 that influences
RT   symbiotic effectiveness consists of three genes encoding an
RT   endoglycanase, a transcriptional regulator and an adenylate cyclase.";
RL   Mol. Gen. Genet. 261:1032-1044(1999).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=1021;
RX   MEDLINE=21396509; PubMed=11481432;
RA   Barnett M.J., Fisher R.F., Jones T., Komp C., Abola A.P.,
RA   Barloy-Hubler F., Bowser L., Capela D., Galibert F., Gouzy J.,
RA   Gurjal M., Hong A., Huizar L., Hyman R.W., Kahn D., Kahn M.L.,
RA   Kalman S., Keating D.H., Palm C., Peck M.C., Surzycki R., Wells D.H.,
RA   Yeh K.-C., Davis R.W., Federspiel N.A., Long S.R.;
RT   "Nucleotide sequence and predicted functions of the entire
RT   Sinorhizobium meliloti pSymA megaplasmid.";
RL   Proc. Natl. Acad. Sci. U.S.A. 98:9883-9888(2001).
CC   -!- FUNCTION: CLEAVES HIGH MOLECULAR WEIGHT SUCCINOGLYCAN TO YIELD LMW
CC       SUCCINOGLYCAN. DYNAMICALLY REGULATES THE MOLECULAR WEIGHT
CC       DISTRIBUTION OF SUCCINOGLYCAN BY CLEAVING NASCENT SUCCINOGLYCAN
CC       ONLY DURING A LIMITED PERIOD AFTER ITS SYNTHESIS, PERHAPS BEFORE
CC       IT UNDERGOES A TIME-DEPENDENT CHANGE IN ITS CONFORMATION OR
CC       AGGREGATION STATE (BY SIMILARITY).
CC   -!- PATHWAY: Exopolysaccharide biosynthesis.
CC   -!- SUBCELLULAR LOCATION: SECRETED. PROBABLY BY A TYPE-III SECRETION
CC       SYSTEM (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO FAMILY 16 OF GLYCOSYL HYDROLASES.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; AJ225896; CAB38101.1; -.
DR   EMBL; AE007273; AAK65522.1; -.
DR   PIR; H95369; H95369.
DR   InterPro; IPR000757; Glyco_hydro_16.
DR   InterPro; IPR001343; Hemlysn_Ca_bind.
DR   Pfam; PF00722; Glyco_hydro_16; 1.
DR   Pfam; PF00353; hemolysinCabind; 3.
DR   PRINTS; PR00313; CABNDNGRPT.
DR   PROSITE; PS01034; GLYCOSYL_HYDROL_F16; FALSE_NEG.
DR   PROSITE; PS00330; HEMOLYSIN_CALCIUM; FALSE_NEG.
KW   Exopolysaccharide synthesis; Glycosidase; Hydrolase; Plasmid;
KW   Complete proteome.
FT   DOMAIN      275    465       CATALYTIC.
FT   ACT_SITE    349    349       NUCLEOPHILE (BY SIMILARITY).
FT   ACT_SITE    354    354       PROTON DONOR (BY SIMILARITY).
FT   CONFLICT     52     52       I -> T (IN REF. 1).
SQ   SEQUENCE   465 AA;  49614 MW;  12CB879AED9E6558 CRC64;

  Query Match              0.9%;  Score 8;  DB 1;  Length 465;
  Best Local Similarity   100.0%;  Pred. No. 13;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        755 NGGAGNDV 762
              ||||||||
Db        123 NGGAGNDV 130


RESULT 18
MAK5_YEAST
ID   MAK5_YEAST     STANDARD;      PRT;   773 AA.
AC   P38112;
DT   01-OCT-1994 (Rel. 30, Created)
DT   01-OCT-1994 (Rel. 30, Last sequence update)
DT   01-NOV-1995 (Rel. 32, Last annotation update)
DE   ATP-dependent RNA helicase MAK5.
GN   MAK5 OR YBR142W OR YBR1119.
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX   NCBI_TaxID=4932;
RN   [1]
RP   SEQUENCE FROM N.A.
```

```
RC   STRAIN=S288c;
RX   MEDLINE=95274325; PubMed=7754712;
RA   Zagulski M., Becam A.-M., Grzybowska E., Lacroute F., Migdalski A.,
RA   Slonimski P.P., Sokolowska B., Herbert C.J.;
RT   "The sequence of 12.5 kb from the right arm of chromosome II predicts
RT   a new N-terminal sequence for the IRA1 protein and reveals two new
RT   genes, one of which is a DEAD-box helicase.";
RL   Yeast 10:1227-1234(1994).
RN   [2]
RP   SEQUENCE OF 770-773 FROM N.A.
RC   STRAIN=S288c;
RA   Entian K.-D., Koetter P., Rose M., Becker J., Grey M., Li Z.,
RA   Niegemann E., Schenk-Groeninger R., Servos J., Wehner E.,
RA   Wolter R., Brendel M., Bauer J., Braun H., Dern K., Duesterhus S.,
RA   Gruenbein R., Hedges D., Kiesau P., Korol S., Krems B., Proft M.,
RA   Siegers K., Baur A., Boles E., Miosga T.,
RA   Schaaff-Gerstenschlaeger I., Zimmermann F.K.;
RL   Submitted (AUG-1994) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: INVOLVED IN MAINTENANCE OF DSRNA KILLER PLASMID.
CC   -!- SIMILARITY: BELONGS TO THE DEAD BOX HELICASE FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; Z36011; CAA85100.1; -.
DR   EMBL; X78937; CAA55539.1; -.
DR   PIR; S46011; S46011.
DR   HSSP; Q58083; 1HV8.
DR   SGD; S0000346; MAK5.
DR   InterPro; IPR001410; DEAD.
DR   InterPro; IPR000629; DEAD_box.
DR   InterPro; IPR001650; Helicase_C.
DR   Pfam; PF00270; DEAD; 1.
DR   Pfam; PF00271; helicase_C; 1.
DR   SMART; SM00487; DEXDc; 1.
DR   SMART; SM00490; HELICc; 1.
DR   PROSITE; PS00039; DEAD_ATP_HELICASE; 1.
KW   Helicase; ATP-binding.
FT   NP_BIND     215    222       ATP (POTENTIAL).
FT   SITE        333    336       DEAD BOX.
SQ   SEQUENCE   773 AA;  87048 MW;  C4FF2FB5B04FFBF9 CRC64;

  Query Match              0.9%;  Score 8;  DB 1;  Length 773;
  Best Local Similarity  100.0%;  Pred. No. 21;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         46 KAADELGI 53
              ||||||||
Db        667 KAADELGI 674


RESULT 19
CHEA_BORBU
ID   CHEA_BORBU     STANDARD;      PRT;   864 AA.
AC   Q44737; P70857; Q44877;
DT   01-NOV-1997 (Rel. 35, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Chemotaxis protein cheA (EC 2.7.3.-).
GN   CHEA OR BB0669.
OS   Borrelia burgdorferi (Lyme disease spirochete).
OC   Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Borrelia.
OX   NCBI_TaxID=139;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=212;
RX   MEDLINE=98438936; PubMed=9765799;
RA   Trueba G.A., Old I.G., Saint-Girons I., Johnson R.C.;
RT   "A cheA cheW operon in Borrelia burgdorferi, the agent of Lyme
RT   disease.";
RL   Res. Microbiol. 148:191-200(1997).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 35210 / B31;
RX   MEDLINE=98065943; PubMed=9403685;
RA   Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,
RA   Lathigra R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M.,
RA   Dougherty B., Tomb J.-F., Fleischmann R.D., Richardson D.,
RA   Peterson J., Kerlavage A.R., Quackenbush J., Salzberg S., Hanson M.,
RA   van Vugt R., Palmer N., Adams M.D., Gocayne J.D., Weidman J.,
RA   Utterback T., Watthey L., McDonald L., Artiach P., Bowman C.,
RA   Garland S., Fujii C., Cotton M.D., Horst K., Roberts K., Hatch B.,
RA   Smith H.O., Venter J.C.;
RT   "Genomic sequence of a Lyme disease spirochaete, Borrelia
RT   burgdorferi.";
RL   Nature 390:580-586(1997).
RN   [3]
RP   SEQUENCE OF 1-30 FROM N.A.
RC   STRAIN=212;
RX   MEDLINE=97144545; PubMed=8990312;
RA   Ge Y., Charon N.W.;
RT   "An unexpected flaA homolog is present and expressed in Borrelia
RT   burgdorferi.";
RL   J. Bacteriol. 179:552-556(1997).
CC   -!- FUNCTION: INVOLVED IN THE TRANSMISSION OF SENSORY SIGNALS FROM
CC       THE CHEMORECEPTORS TO THE FLAGELLAR MOTORS. CHEA IS
CC       AUTOPHOSPHORYLATED; IT CAN TRANSFER ITS PHOSPHATE GROUP TO EITHER
CC       CHEB OR CHEY (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
CC   -!- SIMILARITY: Contains 1 cheW-like domain.
CC   -!- SIMILARITY: Contains 1 histidine kinase domain.
CC   -!- SIMILARITY: Contains 1 HPT domain.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; U28962; AAB96835.1; -.
DR   EMBL; AE001168; AAC67024.1; -.
DR   EMBL; U62900; AAC44771.1; -.
DR   EMBL; X91907; CAA63002.1; -.
DR   PIR; D70183; D70183.
DR   HSSP; Q56310; 1B3Q.
DR   TIGR; BB0669; -.
DR   InterPro; IPR003594; ATPbind_ATPase.
DR   InterPro; IPR004358; Bact_sens_pr_C.
DR   InterPro; IPR002545; CheW.
DR   InterPro; IPR004105; H-kinase_dim.
DR   InterPro; IPR005467; His_kinase.
DR   InterPro; IPR002570; Hpt.
DR   Pfam; PF01584; CheW; 1.
DR   Pfam; PF02895; H-kinase_dim; 1.
DR   Pfam; PF02518; HATPase_c; 1.
DR   Pfam; PF01627; Hpt; 1.
DR   PRINTS; PR00344; BCTRLSENSOR.
DR   ProDom; PD003142; Hpt_N; 1.
DR   SMART; SM00260; CheW; 1.
DR   SMART; SM00387; HATPase_c; 1.
DR   SMART; SM00073; HPT; 1.
DR   PROSITE; PS50851; CHEW; 1.
DR   PROSITE; PS50109; HIS_KIN; 1.
DR   PROSITE; PS50894; HPT; 1.
KW   Sensory transduction; Transferase; Kinase; Phosphorylation;
KW   Chemotaxis; Complete proteome.
FT   DOMAIN        1    108       HPT.
FT   DOMAIN      480    725       HISTIDINE KINASE.
FT   DOMAIN      727    864       CHEW-LIKE.
```

```
FT   MOD_RES       51     51       PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT   CONFLICT      11     11       L -> I (IN REF. 1).
FT   CONFLICT     182    182       S -> H (IN REF. 1).
FT   CONFLICT     187    187       S -> G (IN REF. 1).
FT   CONFLICT     239    239       L -> S (IN REF. 1).
FT   CONFLICT     362    362       A -> S (IN REF. 1).
FT   CONFLICT     559    559       S -> P (IN REF. 1).
SQ   SEQUENCE   864 AA;  98352 MW;  C1111DE0877BE624 CRC64;

  Query Match             0.9%;  Score 8;  DB 1;  Length 864;
  Best Local Similarity   100.0%;  Pred. No. 23;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        869 ELKKLADE 876
              ||||||||
Db        255 ELKKLADE 262


RESULT 20
LKTA_ACTAC
ID   LKTA_ACTAC     STANDARD;      PRT;  1050 AA.
AC   P16462;
DT   01-AUG-1990 (Rel. 15, Created)
DT   01-AUG-1990 (Rel. 15, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Leukotoxin.
GN   LKTA OR LTA.
OS   Actinobacillus actinomycetemcomitans (Haemophilus
OS   actinomycetemcomitans).
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC   Pasteurellaceae; Actinobacillus.
OX   NCBI_TaxID=714;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=JP2;
RX   MEDLINE=89359382; PubMed=2670940;
RA   Lally E.T., Golub E.E., Kieba I.R., Taichman N.S., Rosenbloom J.,
RA   Rosenbloom J.C., Gibson C.W., Demuth D.R.;
RT   "Analysis of the Actinobacillus actinomycetemcomitans leukotoxin
RT   gene. Delineation of unique features and comparison to homologous
RT   toxins.";
RL   J. Biol. Chem. 264:15451-15456(1989).
CC   -!- FUNCTION: One of the virulence factors of A.actinomycetemcomitans
CC       might be a cytotoxin, possibly the membrane-bound hemolysin.
CC   -!- SUBCELLULAR LOCATION: Outer-membrane associated or secreted (By
CC       similarity).
CC   -!- DOMAIN: The Gly-rich region is probably involved in binding
CC       calcium, which is required for target cell-binding or cytolytic
CC       activity.
CC   -!- DOMAIN: The three transmembrane domains are believed to be
CC       involved in pore formation by the cytotoxin (BY SIMILARITY).
CC   -!- PTM: Palmitoylated by lktC. The toxin only becomes active when
CC       modified (By similarity).
CC   -!- MISCELLANEOUS: Its target cell specificity is restricted to human
CC       and some non-human cells of the monomyelocytic lineage.
CC   -!- SIMILARITY: BELONGS TO THE RTX PROKARYOTIC TOXIN FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; M27399; AAA21922.1; -.
DR   InterPro; IPR001343; Hemlysn_Ca_bind.
DR   InterPro; IPR003995; RtxA.
DR   Pfam; PF00353; hemolysinCabind; 7.
DR   Pfam; PF02382; RTX; 1.
DR   PRINTS; PR00313; CABNDNGRPT.
DR   PRINTS; PR01488; RTXTOXINA.
DR   PROSITE; PS00330; HEMOLYSIN_CALCIUM; 5.
KW   Hemolysis; Toxin; Cytolysis; Repeat; Calcium; Transmembrane;
KW   Lipoprotein; Palmitate.
FT   TRANSMEM     339    359       POTENTIAL.
FT   TRANSMEM     408    429       POTENTIAL.
FT   TRANSMEM     477    501       POTENTIAL.
FT   DOMAIN       722    844       14 X REPEATS, GLY-RICH.
FT   REPEAT       722    727       1.
FT   REPEAT       731    736       2.
FT   REPEAT       740    745       3.
FT   REPEAT       749    754       4.
FT   REPEAT       758    763       5.
FT   REPEAT       767    772       6.
FT   REPEAT       776    781       7.
FT   REPEAT       785    790       8.
FT   REPEAT       794    799       9.
FT   REPEAT       803    808       10.
FT   REPEAT       812    817       11.
FT   REPEAT       821    826       12.
FT   REPEAT       830    835       13.
FT   REPEAT       839    844       14.
SQ   SEQUENCE   1050 AA;  114194 MW;  38DF9AA24649F662 CRC64;

  Query Match             0.9%;  Score 8;  DB 1;  Length 1050;
  Best Local Similarity   100.0%;  Pred. No. 27;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        270 AAGLSTTG 277
              ||||||||
Db        295 AAGLSTTG 302


RESULT 21
CYAA_BORBR
ID   CYAA_BORBR     STANDARD;      PRT;  1705 AA.
AC   Q57506; O05179;
DT   01-NOV-1997 (Rel. 35, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Bifunctional hemolysin-adenylate cyclase precursor (Cyclolysin) (ACT)
DE   (AC-HLY) [Contains: Calmodulin-sensitive adenylate cyclase
DE   (EC 4.6.1.1) (ATP pyrophosphate-lyase) (Adenylyl cyclase); Hemolysin].
GN   CYA OR CYAA.
OS   Bordetella bronchiseptica (Alcaligenes bronchisepticus).
OC   Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC   Alcaligenaceae; Bordetella.
OX   NCBI_TaxID=518;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CIP 9.73;
RX   MEDLINE=96009899; PubMed=7557410;
RA   Betsou F., Sismeiro O., Danchin A., Guiso N.;
RT   "Cloning and sequence of the Bordetella bronchiseptica adenylate
RT   cyclase-hemolysin-encoding gene: comparison with the Bordetella
RT   pertussis gene.";
RL   Gene 162:165-166(1995).
RN   [2]
RP   REVISIONS TO 1555-1558.
RA   Danchin A., Boursaux-Eude C.;
RL   Submitted (DEC-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: THIS ADENYLATE CYCLASE BELONGS TO A SPECIAL CLASS OF
CC       BACTERIAL TOXIN. IT CAUSES WHOOPING COUGH BY ACTING ON MAMMALIAN
CC       CELLS BY ELEVATING CAMP-CONCENTRATION AND THUS DISRUPTS NORMAL
CC       CELL FUNCTION.
CC   -!- CATALYTIC ACTIVITY: ATP = 3',5'-cyclic AMP + diphosphate.
CC   -!- ENZYME REGULATION: ACTIVATED BY HOST CALMODULIN.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- DOMAIN: THE GLY-RICH REGION IS PROBABLY INVOLVED IN BINDING
CC       CALCIUM, WHICH IS REQUIRED FOR TARGET CELL-BINDING OR CYTOLYTIC
CC       ACTIVITY (BY SIMILARITY).
CC   -!- PTM: RELEASED IN A PROCESSED FORM.
CC   -!- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE ADENYLYL
CC       CYCLASE CLASS-2 FAMILY.
CC   -!- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE RTX
```

```
CC      PROKARYOTIC TOXIN FAMILY.
CC   -----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----------------------------------------------------------------------
DR   EMBL; Z37112; CAA85481.2; -.
DR   HSSP; P40136; 1K90.
DR   InterPro; IPR005165; Anthrax_toxA.
DR   InterPro; IPR001343; Hemlysn_Ca_bind.
DR   InterPro; IPR003995; RtxA.
DR   Pfam; PF03497; Anthrax_toxA; 1.
DR   Pfam; PF00353; hemolysinCabind; 17.
DR   Pfam; PF02382; RTX; 1.
DR   PRINTS; PR00313; CABNDNGRPT.
DR   PRINTS; PR01488; RTXTOXINA.
DR   PROSITE; PS00330; HEMOLYSIN_CALCIUM; 5.
KW   Lyase; cAMP biosynthesis; ATP-binding; Hemolysis; Toxin; Virulence;
KW   Whooping cough; Calcium-binding; Repeat; Lipoprotein; Palmitate.
FT   CHAIN         1    312       CALMODULIN-SENSITIVE ADENYLATE
FT                                CYCLASE.
FT   CHAIN       313   1705       HEMOLYSIN (BY SIMILARITY TO E.COLI
FT                                HEMOLYSIN HLYA).
FT   DOMAIN        1    399       A, CATALYTIC.
FT   DOMAIN      400    911       B, ALA/GLY-RICH.
FT   DOMAIN      912   1655       C.
FT   DOMAIN     1656   1705       D, ASP/GLY-RICH.
FT   NP_BIND     349    356       ATP (POTENTIAL).
FT   DOMAIN      912   1609       28 X 5 AA REPEATS, GLY-RICH.
FT   REPEAT      912    917       1.
FT   REPEAT     1014   1019       2.
FT   REPEAT     1023   1028       3.
FT   REPEAT     1032   1037       3.
FT   REPEAT     1041   1046       4.
FT   REPEAT     1050   1055       5.
FT   REPEAT     1059   1064       6.
FT   REPEAT     1079   1084       7.
FT   REPEAT     1164   1169       8.
FT   REPEAT     1173   1178       9.
FT   REPEAT     1182   1187       10.
FT   REPEAT     1202   1207       11.
FT   REPEAT     1279   1284       12.
FT   REPEAT     1288   1293       13.
FT   REPEAT     1297   1302       14.
FT   REPEAT     1306   1311       15.
FT   REPEAT     1315   1320       16.
FT   REPEAT     1324   1329       17.
FT   REPEAT     1344   1349       18.
FT   REPEAT     1420   1425       19.
FT   REPEAT     1429   1434       20.
FT   REPEAT     1438   1443       21.
FT   REPEAT     1447   1452       22.
FT   REPEAT     1555   1560       23.
FT   REPEAT     1564   1569       24.
FT   REPEAT     1573   1578       25.
FT   REPEAT     1582   1587       26.
FT   REPEAT     1592   1597       27.
FT   REPEAT     1604   1609       28.
FT   LIPID       859    859       PALMITATE (BY SIMILARITY).
FT   LIPID       982    982       PALMITATE (BY SIMILARITY).
SQ   SEQUENCE   1705 AA;  177153 MW;  C43B30F5886C835A CRC64;

  Query Match             0.9%;  Score 8;  DB 1;  Length 1705;
  Best Local Similarity   100.0%;  Pred. No. 41;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        720 GGAGDDRL 727
              ||||||||
Db       1032 GGAGDDRL 1039
```

```
RESULT 22
RL24_PROVU
ID   RL24_PROVU     STANDARD;      PRT;    18 AA.
AC   P20032;
DT   01-FEB-1991 (Rel. 17, Created)
DT   01-FEB-1991 (Rel. 17, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
DE   50S ribosomal protein L24 (Fragment).
GN   RPLX.
OS   Proteus vulgaris.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Proteus.
OX   NCBI_TaxID=585;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=89125589; PubMed=2464692;
RA   Cerretti D.P., Mattheakis L.C., Kearney K.R., Vu L., Nomura M.;
RT   "Translational regulation of the spc operon in Escherichia coli.
RT   Identification and structural analysis of the target site for S8
RT   repressor protein.";
RL   J. Mol. Biol. 204:309-329(1988).
CC   -!- FUNCTION: THIS PROTEIN IS FOUND IN THE RIBONUCLEOPROTEIN CORE AND
CC       IS INVOLVED IN THE EARLY ASSEMBLY OF THE 50S SUBUNIT. IT IS NOT
CC       INVOLVED IN THE FUNCTIONS OF THE MATURE 50S SUBUNIT.
CC   -!- SIMILARITY: BELONGS TO THE L24P FAMILY OF RIBOSOMAL PROTEINS.
CC   -----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----------------------------------------------------------------------
DR   EMBL; M36264; AAA25661.1; -.
DR   InterPro; IPR005825; Ribosomal_L24_26.
DR   PROSITE; PS01108; RIBOSOMAL_L24; PARTIAL.
KW   Ribosomal protein.
FT   NON_TER       1      1
SQ   SEQUENCE   18 AA;  2202 MW;  43606F53C5CBA57C CRC64;

  Query Match             0.8%;  Score 7;  DB 1;  Length 18;
  Best Local Similarity   100.0%;  Pred. No. 8.5;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        481 FEDGKKV 487
              |||||||
Db          1 FEDGKKV 7


RESULT 23
CECA_BOMMO
ID   CECA_BOMMO     STANDARD;      PRT;    63 AA.
AC   Q27239;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Cecropin A precursor.
GN   CECA.
OS   Bombyx mori (Silk moth).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia; Bombycoidea;
OC   Bombycidae; Bombyx.
OX   NCBI_TaxID=7091;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C108; TISSUE=Larval fat body;
RX   MEDLINE=94369101; PubMed=7765280;
RA   Yamano Y., Matsumoto M., Inoue K., Kawabata T., Morishima I.;
RT   "Cloning of cDNAs for cecropins A and B, and expression of the genes
RT   in the silkworm, Bombyx mori.";
```

```
RL   Biosci. Biotechnol. Biochem. 58:1476-1478(1994).
CC   -!- FUNCTION: CECROPINS HAVE LYTIC AND ANTIBACTERIAL ACTIVITY AGAINST
CC       SEVERAL GRAM-POSITIVE AND GRAM-NEGATIVE BACTERIA.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: Highest expression in fat body and hemocytes.
CC       Is also expressed in Malpighian tubules and to a much lesser
CC       extent in midgut. Not present in silk gland.
CC   -!- SIMILARITY: BELONGS TO THE CECROPIN FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; D17394; BAA04217.1; -.
DR   EMBL; S74376; AAC60515.1; -.
DR   PIR; JC2295; CKMTA.
DR   InterPro; IPR000875; Cecropin.
DR   InterPro; IPR003254; IIP_cecropin.
DR   Pfam; PF00272; cecropin; 1.
DR   ProDom; PD003996; IIP_cecropin; 1.
DR   PROSITE; PS00268; CECROPIN; 1.
KW   Insect immunity; Antibiotic; Hemolymph; Amidation; Multigene family;
KW   Signal.
FT   SIGNAL        1     22       POTENTIAL.
FT   PROPEP       23     26       BY SIMILARITY.
FT   CHAIN        27     61       CECROPIN A.
FT   MOD_RES      61     61       AMIDATION (G-62 PROVIDE AMIDE GROUP)
FT                                (BY SIMILARITY).
SQ   SEQUENCE   63 AA;  6762 MW;  6A1C39975516D86A CRC64;

  Query Match             0.8%;  Score 7;  DB 1;  Length 63;
  Best Local Similarity   100.0%;  Pred. No. 25;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        347 ALGAVSA 353
              |||||||
Db         16 ALGAVSA 22


RESULT 24
SCX1_TITSE
ID   SCX1_TITSE     STANDARD;      PRT;    80 AA.
AC   P01496; P91788;
DT   21-JUL-1986 (Rel. 01, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Toxin IV-5 precursor (Tityustoxin) (Fragment).
OS   Tityus serrulatus (Brazilian scorpion).
OC   Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;
OC   Buthoidea; Buthidae; Tityus.
OX   NCBI_TaxID=6887;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96243503; PubMed=8711758;
RA   Corona M., Zurita M., Possani L.D., Becerril B.;
RT   "Cloning and characterization of the genomic region encoding toxin
RT   IV-5 from the scorpion Tityus serrulatus Lutz and Mello.";
RL   Toxicon 34:251-256(1996).
RN   [2]
RP   SEQUENCE OF 14-75.
RC   TISSUE=Venom;
RX   MEDLINE=91131623; PubMed=1993690;
RA   Possani L.D., Martin B.M., Fletcher M.D., Fletcher P.L. Jr.;
RT   "Discharge effect on pancreatic exocrine secretion produced by toxins
RT   purified from Tityus serrulatus scorpion venom.";
RL   J. Biol. Chem. 266:3178-3185(1991).
RN   [3]
RP   SEQUENCE OF 14-43.
RC   TISSUE=Venom;
RA   Possani L.D., Martin B.M., Mochca-Morales J., Svendsen I.;
RT   "N-terminal sequence of toxin IV-5 from the venom of the scorpion
RT   Tytius serrulatus.";
RL   Toxicon 20:75-76(1982).
CC   -!- FUNCTION: Binds to sodium channels and inhibits the inactivation
CC       of the activated channels, thereby blocking neuronal transmission.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: Expressed by the venom gland.
CC   -!- SIMILARITY: BELONGS TO THE ALPHA/BETA-SCORPION TOXIN FAMILY.
CC       BETA-TOXIN SUBFAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; S82286; AAB37719.2; -.
DR   HSSP; P46066; 1NRB.
DR   InterPro; IPR003614; Knot1.
DR   InterPro; IPR002061; Scorpion_toxinL.
DR   Pfam; PF00537; toxin_3; 1.
DR   ProDom; PD000908; Scorpion_toxinL; 1.
DR   SMART; SM00505; Knot1; 1.
KW   Toxin; Neurotoxin; Ionic channel inhibitor; Sodium channel inhibitor;
KW   Amidation; Signal.
FT   NON_TER       1      1
FT   SIGNAL       <1     13
FT   CHAIN        14     77       TOXIN IV-5.
FT   DISULFID     25     75       BY SIMILARITY.
FT   DISULFID     29     51       BY SIMILARITY.
FT   DISULFID     37     58       BY SIMILARITY.
FT   DISULFID     41     60       BY SIMILARITY.
FT   MOD_RES      77     77       AMIDATION (G-78 PROVIDE AMIDE GROUP)
FT                                (POTENTIAL).
FT   CONFLICT     70     73       KTNG -> GST (IN REF. 2).
SQ   SEQUENCE   80 AA;  9043 MW;  974F46346A435FFD CRC64;

  Query Match             0.8%;  Score 7;  DB 1;  Length 80;
  Best Local Similarity   100.0%;  Pred. No. 31;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        521 LLTAGTE 527
              |||||||
Db          6 LLTAGTE 12


RESULT 25
SCX4_TITSE
ID   SCX4_TITSE     STANDARD;      PRT;    80 AA.
AC   P45659;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Toxins IV/III/V precursor (TsTX-IV) (Tityustoxin IV) (Ts IV)
DE   (Fragment).
OS   Tityus serrulatus (Brazilian scorpion).
OC   Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;
OC   Buthoidea; Buthidae; Tityus.
OX   NCBI_TaxID=6887;
RN   [1]
RP   SEQUENCE FROM N.A., AND SEQUENCE OF 14-43.
RC   TISSUE=Venom;
RX   MEDLINE=94192827; PubMed=8143874;
RA   Martin-Eauclaire M.-F., Ceard B., Ribeiro A.M., Diniz C.R., Rochat H.,
RA   Bougis P.E.;
RT   "Biochemical, pharmacological and genomic characterisation of Ts IV,
RT   an alpha-toxin from the venom of the South American scorpion Tityus
RT   serrulatus.";
RL   FEBS Lett. 342:181-184(1994).
CC   -!- FUNCTION: Binds to sodium channels and inhibits the inactivation
```

```
CC       of the activated channels, thereby blocking neuronal transmission.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: Expressed by the venom gland.
CC   -!- SIMILARITY: BELONGS TO THE ALPHA/BETA-SCORPION TOXIN FAMILY.
CC       ALPHA-TOXIN SUBFAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; S69808; AAB30413.1; -.
DR   PIR; S43674; NTSR4T.
DR   HSSP; P46066; 1NRB.
DR   InterPro; IPR003614; Knot1.
DR   InterPro; IPR002061; Scorpion_toxinL.
DR   Pfam; PF00537; toxin_3; 1.
DR   ProDom; PD000908; Scorpion_toxinL; 1.
DR   SMART; SM00505; Knot1; 1.
KW   Toxin; Neurotoxin; Ionic channel inhibitor; Sodium channel inhibitor;
KW   Signal; Amidation.
FT   NON_TER       1      1
FT   SIGNAL       <1     13
FT   CHAIN        14     77       TOXIN IV.
FT   CHAIN        14     79       TOXIN V (POTENTIAL).
FT   CHAIN        14     75       TOXIN III (POTENTIAL).
FT   DISULFID     25     75       BY SIMILARITY.
FT   DISULFID     29     51       BY SIMILARITY.
FT   DISULFID     37     58       BY SIMILARITY.
FT   DISULFID     41     60       BY SIMILARITY.
FT   MOD_RES      77     77       AMIDATION (G-78 PROVIDE AMIDE GROUP)
FT                                (POTENTIAL).
SQ   SEQUENCE   80 AA;  9025 MW;  FDE437EEC1335FFC CRC64;

  Query Match              0.8%;  Score 7;  DB 1;  Length 80;
  Best Local Similarity   100.0%;  Pred. No. 31;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        521 LLTAGTE 527
              |||||||
Db          6 LLTAGTE 12


RESULT 26
TRAC_RHISN
ID   TRAC_RHISN     STANDARD;      PRT;   102 AA.
AC   P55419;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   01-NOV-1997 (Rel. 35, Last annotation update)
DE   Probable conjugal transfer protein traC.
GN   TRAC OR Y4DT.
OS   Rhizobium sp. (strain NGR234).
OG   Plasmid sym pNGR234a.
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC   Rhizobiaceae; Rhizobium/Agrobacterium group; Rhizobium.
OX   NCBI_TaxID=394;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=97305956; PubMed=9163424;
RA   Freiberg C.A., Fellay R., Bairoch A., Broughton W.J., Rosenthal A.,
RA   Perret X.;
RT   "Molecular basis of symbiosis between Rhizobium and legumes.";
RL   Nature 387:394-401(1997).
CC   -!- SIMILARITY: STRONG, TO A.TUMEFACIENS TI PLASMID TRAC.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; AE000070; AAB92440.1; -.
KW   Conjugation; Plasmid.
SQ   SEQUENCE   102 AA;  10205 MW;  AA378F93DD06C78B CRC64;

  Query Match              0.8%;  Score 7;  DB 1;  Length 102;
  Best Local Similarity   100.0%;  Pred. No. 38;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        356 SAAAVGS 362
              |||||||
Db         86 SAAAVGS 92


RESULT 27
RL24_ECOLI
ID   RL24_ECOLI     STANDARD;      PRT;   103 AA.
AC   P02425; P37438;
DT   21-JUL-1986 (Rel. 01, Created)
DT   21-JUL-1986 (Rel. 01, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   50S ribosomal protein L24.
GN   RPLX OR B3309 OR Z4679 OR ECS4174 OR STM3429.
OS   Escherichia coli,
OS   Escherichia coli O157:H7, and
OS   Salmonella typhimurium.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Escherichia.
OX   NCBI_TaxID=562, 83334, 602;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   SPECIES=E.coli; STRAIN=K12;
RX   MEDLINE=83220807; PubMed=6222285;
RA   Ceretti D.P., Dean D., Davis G.R., Bedwell D.M., Nomura M.;
RT   "The spc ribosomal protein operon of Escherichia coli: sequence and
RT   cotranscription of the ribosomal protein genes and a protein export
RT   gene.";
RL   Nucleic Acids Res. 11:2599-2616(1983).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   SPECIES=E.coli; STRAIN=K12 / MG1655;
RX   MEDLINE=97426617; PubMed=9278503;
RA   Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA   Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA   Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA   Mau B., Shao Y.;
RT   "The complete genome sequence of Escherichia coli K-12.";
RL   Science 277:1453-1474(1997).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   SPECIES=E.coli; STRAIN=O157:H7 / EDL933 / ATCC 700927;
RX   MEDLINE=21074935; PubMed=11206551;
RA   Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA   Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA   Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA   Grotbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,
RA   Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA   Welch R.A., Blattner F.R.;
RT   "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
RL   Nature 409:529-533(2001).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   SPECIES=E.coli; STRAIN=O157:H7 / RIMD 0509952;
RX   MEDLINE=21156231; PubMed=11258796;
RA   Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA   Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA   Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
RA   Kuhara S., Shiba T., Hattori M., Shinagawa H.;
RT   "Complete genome sequence of enterohemorrhagic Escherichia coli
RT   O157:H7 and genomic comparison with a laboratory strain K-12.";
```

```
RL   DNA Res. 8:11-22(2001).
RN   [5]
RP   SEQUENCE.
RC   SPECIES=E.coli; STRAIN=K12;
RX   MEDLINE=80092112; PubMed=391595;
RA   Wittmann-Liebold B.;
RT   "Primary structure of protein L24 from the Escherichia coli
RT   ribosome.";
RL   FEBS Lett. 108:75-80(1979).
RN   [6]
RP   SEQUENCE OF 85-103 FROM N.A.
RC   SPECIES=E.coli;
RX   MEDLINE=81199003; PubMed=6262737;
RA   Olins P.O., Nomura M.;
RT   "Translational regulation by ribosomal protein S8 in Escherichia
RT   coli: structural homology between rRNA binding site and feedback
RT   target on mRNA.";
RL   Nucleic Acids Res. 9:1757-1764(1981).
RN   [7]
RP   SEQUENCE FROM N.A.
RC   SPECIES=S.typhimurium; STRAIN=LT2 / SGSC1412 / ATCC 700720;
RX   MEDLINE=21534948; PubMed=11677609;
RA   McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA   Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
RA   Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA   Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
RA   Waterston R., Wilson R.K.;
RT   "Complete genome sequence of Salmonella enterica serovar Typhimurium
RT   LT2.";
RL   Nature 413:852-856(2001).
RN   [8]
RP   SEQUENCE OF 86-103 FROM N.A.
RC   SPECIES=S.typhimurium;
RX   MEDLINE=89125589; PubMed=2464692;
RA   Cerretti D.P., Mattheakis L.C., Kearney K.R., Vu L., Nomura M.;
RT   "Translational regulation of the spc operon in Escherichia coli.
RT   Identification and structural analysis of the target site for S8
RT   repressor protein.";
RL   J. Mol. Biol. 204:309-329(1988).
RN   [9]
RP   MASS SPECTROMETRY.
RC   SPECIES=E.coli; STRAIN=K12 / ATCC 25404;
RX   MEDLINE=99196679; PubMed=10094780;
RA   Arnold R.J., Reilly J.P.;
RT   "Observation of Escherichia coli ribosomal proteins and their
RT   posttranslational modifications by mass spectrometry.";
RL   Anal. Biochem. 269:105-112(1999).
CC   -!- FUNCTION: THIS PROTEIN IS FOUND IN THE RIBONUCLEOPROTEIN CORE AND
CC       IS INVOLVED IN THE EARLY ASSEMBLY OF THE 50S SUBUNIT. IT IS NOT
CC       INVOLVED IN THE FUNCTIONS OF THE MATURE 50S SUBUNIT.
CC   -!- MASS SPECTROMETRY: MW=11186.5; METHOD=MALDI.
CC   -!- SIMILARITY: BELONGS TO THE L24P FAMILY OF RIBOSOMAL PROTEINS.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; X01563; CAA25716.1; -.
DR   EMBL; U18997; AAA58106.1; -.
DR   EMBL; AE000408; AAC76334.1; -.
DR   EMBL; AE005556; AAG58430.1; -.
DR   EMBL; AP002564; BAB37597.1; -.
DR   EMBL; M10195; AAA24050.1; -.
DR   EMBL; AE008857; AAL22292.1; -.
DR   EMBL; M36266; AAA27227.1; -.
DR   PIR; B85996; B85996.
DR   PIR; F91150; F91150.
DR   PIR; H65123; R5EC24.
DR   EcoGene; EG10884; rplX.
DR   StyGene; SG10366; rplX.
DR   InterPro; IPR005824; KOW.
DR   InterPro; IPR006646; KOW_sub.
DR   InterPro; IPR003256; Ribosomal_L24.
DR   InterPro; IPR005825; Ribosomal_L24_26.
DR   Pfam; PF00467; KOW; 1.
DR   ProDom; PD001677; Ribosomal_L24; 1.
DR   SMART; SM00739; KOW; 1.
DR   TIGRFAMs; TIGR01079; rplX_bact; 1.
DR   PROSITE; PS01108; RIBOSOMAL_L24; 1.
KW   Ribosomal protein; Complete proteome.
FT   INIT_MET      0       0
SQ   SEQUENCE   103 AA;  11185 MW;  9F4CB2EFF66FAE95 CRC64;

  Query Match             0.8%;  Score 7;  DB 1;  Length 103;
  Best Local Similarity   100.0%;  Pred. No. 38;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy.        481 FEDGKKV 487
               |||||||
Db          86 FEDGKKV 92


RESULT 28
RL24_HAEIN
ID   RL24_HAEIN     STANDARD;      PRT;   103 AA.
AC   P44362;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   50S ribosomal protein L24.
GN   RPLX OR RPL24 OR HI0789.
OS   Haemophilus influenzae.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC   Pasteurellaceae; Haemophilus.
OX   NCBI_TaxID=727;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Rd / KW20 / ATCC 51907;
RX   MEDLINE=95350630; PubMed=7542800;
RA   Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA   Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA   McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA   Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
RA   Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA   Utterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RA   Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
RA   Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA   Venter J.C.;
RT   "Whole-genome random sequencing and assembly of Haemophilus influenzae
RT   Rd.";
RL   Science 269:496-512(1995).
CC   -!- FUNCTION: THIS PROTEIN IS FOUND IN THE RIBONUCLEOPROTEIN CORE AND
CC       IS INVOLVED IN THE EARLY ASSEMBLY OF THE 50S SUBUNIT. IT IS NOT
CC       INVOLVED IN THE FUNCTIONS OF THE MATURE 50S SUBUNIT
CC       (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE L24P FAMILY OF RIBOSOMAL PROTEINS.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; U32762; AAC22447.1; -.
DR   PIR; G64093; G64093.
DR   TIGR; HI0789; -.
DR   InterPro; IPR005824; KOW.
DR   InterPro; IPR006646; KOW_sub.
DR   InterPro; IPR003256; Ribosomal_L24.
DR   InterPro; IPR005825; Ribosomal_L24_26.
```

```
DR   Pfam; PF00467; KOW; 1.
DR   ProDom; PD001677; Ribosomal_L24; 1.
DR   SMART; SM00739; KOW; 1.
DR   TIGRFAMs; TIGR01079; rplX_bact; 1.
DR   PROSITE; PS01108; RIBOSOMAL_L24; 1.
KW   Ribosomal protein; Complete proteome.
SQ   SEQUENCE   103 AA;  11285 MW;  3AD4903377702F19 CRC64;

  Query Match             0.8%;  Score 7;  DB 1;  Length 103;
  Best Local Similarity   100.0%;  Pred. No. 38;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        481 FEDGKKV 487
              |||||||
Db         87 FEDGKKV 93


RESULT 29
RL24_PASMU
ID   RL24_PASMU     STANDARD;      PRT;   103 AA.
AC   Q9CL41;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   50S ribosomal protein L24.
GN   RPLX OR RPL24 OR PM1404.
OS   Pasteurella multocida.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC   Pasteurellaceae; Pasteurella.
OX   NCBI_TaxID=747;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Pm70;
RX   MEDLINE=21145866; PubMed=11248100;
RA   May B.J., Zhang Q., Li L.L., Paustian M.L., Whittam T.S., Kapur V.;
RT   "Complete genomic sequence of Pasteurella multocida Pm70.";
RL   Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).
CC   -!- FUNCTION: THIS PROTEIN IS FOUND IN THE RIBONUCLEOPROTEIN CORE AND
CC       IS INVOLVED IN THE EARLY ASSEMBLY OF THE 50S SUBUNIT. IT IS NOT
CC       INVOLVED IN THE FUNCTIONS OF THE MATURE 50S SUBUNIT
CC       (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE L24P FAMILY OF RIBOSOMAL PROTEINS.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; AE006177; AAK03488.1; -.
DR   InterPro; IPR005824; KOW.
DR   InterPro; IPR006646; KOW_sub.
DR   InterPro; IPR003256; Ribosomal_L24.
DR   InterPro; IPR005825; Ribosomal_L24_26.
DR   Pfam; PF00467; KOW; 1.
DR   ProDom; PD001677; Ribosomal_L24; 1.
DR   SMART; SM00739; KOW; 1.
DR   TIGRFAMs; TIGR01079; rplX_bact; 1.
DR   PROSITE; PS01108; RIBOSOMAL_L24; 1.
KW   Ribosomal protein; Complete proteome.
SQ   SEQUENCE   103 AA;  11239 MW;  0A325EF73E604850 CRC64;

  Query Match             0.8%;  Score 7;  DB 1;  Length 103;
  Best Local Similarity   100.0%;  Pred. No. 38;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        481 FEDGKKV 487
              |||||||
Db         87 FEDGKKV 93


RESULT 30
RL24_BUCAK
ID   RL24_BUCAK     STANDARD;      PRT;   104 AA.
AC   P46177;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   50S ribosomal protein L24.
GN   RPLX.
OS   Buchnera aphidicola (subsp. Acyrthosiphon kondoi) (Acyrthosiphon
OS   kondoi symbiotic bacterium).
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Buchnera.
OX   NCBI_TaxID=42474;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Kurashiki;
RX   MEDLINE=96051390; PubMed=7584036;
RA   Abe R., Yamashita A., Isono K.;
RT   "Cloning and characterization of the ribosomal protein genes in the
RT   spc operon of a prokaryotic endosymbiont of the pea aphid,
RT   Acyrthosiphon kondoi.";
RL   DNA Res. 1:103-114(1994).
CC   -!- FUNCTION: THIS PROTEIN IS FOUND IN THE RIBONUCLEOPROTEIN CORE AND
CC       IS INVOLVED IN THE EARLY ASSEMBLY OF THE 50S SUBUNIT. IT IS NOT
CC       INVOLVED IN THE FUNCTIONS OF THE MATURE 50S SUBUNIT
CC       (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE L24P FAMILY OF RIBOSOMAL PROTEINS.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; D31786; BAA06586.1; -.
DR   InterPro; IPR005824; KOW.
DR   InterPro; IPR006646; KOW_sub.
DR   InterPro; IPR003256; Ribosomal_L24.
DR   InterPro; IPR005825; Ribosomal_L24_26.
DR   Pfam; PF00467; KOW; 1.
DR   ProDom; PD001677; Ribosomal_L24; 1.
DR   SMART; SM00739; KOW; 1.
DR   TIGRFAMs; TIGR01079; rplX_bact; 1.
DR   PROSITE; PS01108; RIBOSOMAL_L24; 1.
KW   Ribosomal protein.
SQ   SEQUENCE   104 AA;  11336 MW;  F3553FE6418BF47C CRC64;

  Query Match             0.8%;  Score 7;  DB 1;  Length 104;
  Best Local Similarity   100.0%;  Pred. No. 38;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        481 FEDGKKV 487
              |||||||
Db         87 FEDGKKV 93


RESULT 31
NU4M_CAICR
ID   NU4M_CAICR     STANDARD;      PRT;   111 AA.
AC   Q34076;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   NADH-ubiquinone oxidoreductase chain 4 (EC 1.6.5.3) (Fragment).
GN   MTND4 OR ND4.
OS   Caiman crocodilus (Spectacled caiman) (Caiman sclerops).
OG   Mitochondrion.
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Crocodylidae; Alligatorinae; Caiman.
OX   NCBI_TaxID=8499;
```

```
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96073446; PubMed=7476123;
RA   Kumazawa Y., Nishida M.;
RT   "Variations in mitochondrial tRNA gene organization of reptiles as
RT   phylogenetic markers.";
RL   Mol. Biol. Evol. 12:759-772(1995).
CC   -!- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; D38190; BAA07383.1; -.
DR   InterPro; IPR001750; Oxidored_q1.
DR   Pfam; PF00361; oxidored_q1; 1.
KW   Oxidoreductase; NAD; Ubiquinone; Mitochondrion.
FT   NON_TER       1      1
SQ   SEQUENCE   111 AA;  12070 MW;  D211FBBA07A42D83 CRC64;

  Query Match             0.8%;  Score 7;  DB 1;  Length 111;
  Best Local Similarity   100.0%;  Pred. No. 40;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        518 TSPLLTA 524
              |||||||
Db          5 TSPLLTA 11


RESULT 32
RL7_NEILA
ID   RL7_NEILA      STANDARD;      PRT;   122 AA.
AC   Q9ETV2;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   50S ribosomal protein L7/L12.
GN   RPLL.
OS   Neisseria lactamica, and
OS   Neisseria sicca.
OC   Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC   Neisseriaceae; Neisseria.
OX   NCBI_TaxID=486, 490;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   SPECIES=N.lactamica, and N.sicca; STRAIN=NRL 8828, and NRL 30016;
RA   Spence J.M., Clark V.L.;
RT   "Alterations in protein profiles associated with induction of the
RT   contact-induced enhanced invasion phenotype of Neisseria
RT   gonorrhoeae.";
RL   Submitted (OCT-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: SEEMS TO BE THE BINDING SITE FOR SEVERAL OF THE FACTORS
CC       INVOLVED IN PROTEIN SYNTHESIS AND APPEARS TO BE ESSENTIAL FOR
CC       ACCURATE TRANSLATION (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE L12P FAMILY OF RIBOSOMAL PROTEINS.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; AF312972; AAG34164.1; -.
DR   EMBL; AF312974; AAG34166.1; -.
DR   HSSP; P02392; 1CTF.
DR   HAMAP; MF_00368; -; 1.
DR   InterPro; IPR000206; Ribosomal_L12.
DR   Pfam; PF00542; Ribosomal_L12; 1.
DR   ProDom; PD001326; Ribosomal_L12; 1.
DR   TIGRFAMs; TIGR00855; L12; 1.
KW   Ribosomal protein.
SQ   SEQUENCE   122 AA;  12567 MW;  7A87249AAC0E2D4D CRC64;

  Query Match             0.8%;  Score 7;  DB 1;  Length 122;
  Best Local Similarity   100.0%;  Pred. No. 44;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        354 GVSAAAV 360
              |||||||
Db         32 GVSAAAV 38


RESULT 33
RL7_NEIMA
ID   RL7_NEIMA      STANDARD;      PRT;   122 AA.
AC   P80716;
DT   01-OCT-1996 (Rel. 34, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   50S ribosomal protein L7/L12.
GN   RPLL OR NMA0143 OR NMB0131.
OS   Neisseria meningitidis (serogroup A), and
OS   Neisseria meningitidis (serogroup B).
OC   Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC   Neisseriaceae; Neisseria.
OX   NCBI_TaxID=65699, 491;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Z2491 / Serogroup A / Serotype 4A;
RX   MEDLINE=20222556; PubMed=10761919;
RA   Parkhill J., Achtman M., James K.D., Bentley S.D., Churcher C.,
RA   Klee S.R., Morelli G., Basham D., Brown D., Chillingworth T.,
RA   Davies R.M., Davis P., Devlin K., Feltwell T., Hamlin N., Holroyd S.,
RA   Jagels K., Leather S., Moule S., Mungall K., Quail M.A.,
RA   Rajandream M.A., Rutherford K.M., Simmonds M., Skelton J.,
RA   Whitehead S., Spratt B.G., Barrell B.G.;
RT   "Complete DNA sequence of a serogroup A strain of Neisseria
RT   meningitidis Z2491.";
RL   Nature 404:502-506(2000).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MC58 / Serogroup B;
RX   MEDLINE=20175755; PubMed=10710307;
RA   Tettelin H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E.,
RA   Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F., Dodson R.J.,
RA   Nelson W.C., Gwinn M.L., DeBoy R., Peterson J.D., Hickey E.K.,
RA   Haft D.H., Salzberg S.L., White O., Fleischmann R.D., Dougherty B.A.,
RA   Mason T., Ciecko A., Parksey D.S., Blair E., Cittone H., Clark E.B.,
RA   Cotton M.D., Utterback T.R., Khouri H., Qin H., Vamathevan J.,
RA   Gill J., Scarlato V., Masignani V., Pizza M., Grandi G., Sun L.,
RA   Smith H.O., Fraser C.M., Moxon E.R., Rappuoli R., Venter J.C.;
RT   "Complete genome sequence of Neisseria meningitidis serogroup B strain
RT   MC58.";
RL   Science 287:1809-1815(2000).
RN   [3]
RP   SEQUENCE OF 1-20.
RC   STRAIN=44/76 / Serogroup B / Serotype 15 / Subtype 16;
RX   MEDLINE=97177772; PubMed=9025278;
RA   Kolberg J., Hoiby E.A., Lopez R., Sletten K.;
RT   "Monoclonal antibodies against Streptococcus pneumoniae detect
RT   epitopes on eubacterial ribosomal proteins L7/L12 and on
RT   streptococcal elongation factor Ts.";
RL   Microbiology 143:55-61(1997).
CC   -!- FUNCTION: SEEMS TO BE THE BINDING SITE FOR SEVERAL OF THE FACTORS
CC       INVOLVED IN PROTEIN SYNTHESIS AND APPEARS TO BE ESSENTIAL FOR
CC       ACCURATE TRANSLATION (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE L12P FAMILY OF RIBOSOMAL PROTEINS.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
```

```
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----------------------------------------------------------------------
DR   EMBL; AL162752; CAB83458.1; -.
DR   EMBL; AE002371; AAF40590.1; -.
DR   PIR; H81235; H81235.
DR   HSSP; P02392; 1CTF.
DR   TIGR; NMB0131; -.
DR   HAMAP; MF_00368; -; 1.
DR   InterPro; IPR000206; Ribosomal_L12.
DR   Pfam; PF00542; Ribosomal_L12; 1.
DR   ProDom; PD001326; Ribosomal_L12; 1.
DR   TIGRFAMs; TIGR00855; L12; 1.
KW   Ribosomal protein; Complete proteome.
FT   INIT_MET      0      0
SQ   SEQUENCE   122 AA;  12491 MW;  3C962C3A11B6F535 CRC64;

  Query Match             0.8%;  Score 7;  DB 1;  Length 122;
  Best Local Similarity   100.0%;  Pred. No. 44;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        354 GVSAAAV 360
              |||||||
Db         31 GVSAAAV 37


RESULT 34
RL7_YERPE
ID   RL7_YERPE      STANDARD;      PRT;   122 AA.
AC   Q8ZAP4;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   50S ribosomal protein L7/L12.
GN   RPLL OR YPO3748 OR Y0483.
OS   Yersinia pestis.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Yersinia.
OX   NCBI_TaxID=632;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CO-92 / Biovar Orientalis;
RX   MEDLINE=21470413; PubMed=11586360;
RA   Parkhill J., Wren B.W., Thomson N.R., Titball R.W., Holden M.T.G.,
RA   Prentice M.B., Sebaihia M., James K.D., Churcher C., Mungall K.L.,
RA   Baker S., Basham D., Bentley S.D., Brooks K., Cerdeno-Tarraga A.M.,
RA   Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,
RA   Feltwell T., Hamlin N., Holroyd S., Jagels K., Karlyshev A.V.,
RA   Leather S., Moule S., Oyston P.C.F., Quail M., Rutherford K.,
RA   Simmonds M., Skelton J., Stevens K., Whitehead S., Barrell B.G.;
RT   "Genome sequence of Yersinia pestis, the causative agent of plague.";
RL   Nature 413:523-527(2001).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=KIM5 / Biovar Mediaevalis;
RX   MEDLINE=22137863; PubMed=12142430;
RA   Deng W., Burland V., Plunkett G. III, Boutin A., Mayhew G.F., Liss P.,
RA   Perna N.T., Rose D.J., Mau B., Zhou S., Schwartz D.C.,
RA   Fetherston J.D., Lindler L.E., Brubaker R.R., Plano G.V.,
RA   Straley S.C., McDonough K.A., Nilles M.L., Matson J.S., Blattner F.R.,
RA   Perry R.D.;
RT   "Genome sequence of Yersinia pestis KIM.";
RL   J. Bacteriol. 184:4601-4611(2002).
CC   -!- FUNCTION: Seems to be the binding site for several of the factors
CC       involved in protein synthesis and appears to be essential for
CC       accurate translation (By similarity).
CC   -!- SIMILARITY: BELONGS TO THE L12P FAMILY OF RIBOSOMAL PROTEINS.
CC   -----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----------------------------------------------------------------------
DR   EMBL; AJ414158; CAC93216.1; -.
DR   EMBL; AE013649; AAM84072.1; -.
DR   PIR; AD0456; AD0456.
DR   HAMAP; MF_00368; -; 1.
DR   InterPro; IPR000206; Ribosomal_L12.
DR   Pfam; PF00542; Ribosomal_L12; 1.
DR   ProDom; PD001326; Ribosomal_L12; 1.
DR   TIGRFAMs; TIGR00855; L12; 1.
KW   Ribosomal protein; Complete proteome.
SQ   SEQUENCE   122 AA;  12530 MW;  40A0C74D6B77D69D CRC64;

  Query Match             0.8%;  Score 7;  DB 1;  Length 122;
  Best Local Similarity   100.0%;  Pred. No. 44;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        354 GVSAAAV 360
              |||||||
Db         33 GVSAAAV 39


RESULT 35
RL7_NEIPE
ID   RL7_NEIPE      STANDARD;      PRT;   123 AA.
AC   Q9F5M1;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   50S ribosomal protein L7/L12.
GN   RPLL.
OS   Neisseria perflava.
OC   Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC   Neisseriaceae; Neisseria.
OX   NCBI_TaxID=33053;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NRL 30015;
RA   Spence J.M., Clark V.L.;
RT   "Alterations in protein profiles associated with induction of the
RT   contact-induced enhanced invasion phenotype of Neisseria
RT   gonorrhoeae.";
RL   Submitted (OCT-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: SEEMS TO BE THE BINDING SITE FOR SEVERAL OF THE FACTORS
CC       INVOLVED IN PROTEIN SYNTHESIS AND APPEARS TO BE ESSENTIAL FOR
CC       ACCURATE TRANSLATION (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE L12P FAMILY OF RIBOSOMAL PROTEINS.
CC   -----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----------------------------------------------------------------------
DR   EMBL; AF312973; AAG34165.1; -.
DR   HSSP; P02392; 1CTF.
DR   HAMAP; MF_00368; -; 1.
DR   InterPro; IPR000206; Ribosomal_L12.
DR   Pfam; PF00542; Ribosomal_L12; 1.
DR   ProDom; PD001326; Ribosomal_L12; 1.
DR   TIGRFAMs; TIGR00855; L12; 1.
KW   Ribosomal protein.
SQ   SEQUENCE   123 AA;  12578 MW;  5A6257CB0DE3A435 CRC64;

  Query Match             0.8%;  Score 7;  DB 1;  Length 123;
  Best Local Similarity   100.0%;  Pred. No. 44;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy        354 GVSAAAV 360
              |||||||
Db         32 GVSAAAV 38


RESULT 36
RL7_RALSO
ID   RL7_RALSO      STANDARD;      PRT;   124 AA.
AC   Q8XUZ7;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   50S ribosomal protein L7/L12.
GN   RPLL OR RSC3035 OR RS04722.
OS   Ralstonia solanacearum (Pseudomonas solanacearum).
OC   Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC   Burkholderiaceae; Ralstonia.
OX   NCBI_TaxID=305;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=GMI1000;
RX   MEDLINE=21681879; PubMed=11823852;
RA   Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,
RA   Arlat M., Billault A., Brottier P., Camus J.C., Cattolico L.,
RA   Chandler M., Choisne N., Claudel-Renard C., Cunnac S., Demange N.,
RA   Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schiex T.,
RA   Siguier P., Thebault P., Whalen M., Wincker P., Levy M.,
RA   Weissenbach J., Boucher C.A.;
RT   "Genome sequence of the plant pathogen Ralstonia solanacearum.";
RL   Nature 415:497-502(2002).
CC   -!- FUNCTION: Seems to be the binding site for several of the factors
CC       involved in protein synthesis and appears to be essential for
CC       accurate translation (By similarity).
CC   -!- SIMILARITY: BELONGS TO THE L12P FAMILY OF RIBOSOMAL PROTEINS.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; AL646073; CAD16744.1; -.
DR   HAMAP; MF_00368; -; 1.
DR   InterPro; IPR001813; 60s_ribosomal.
DR   InterPro; IPR000206; Ribosomal_L12.
DR   Pfam; PF00428; 60s_ribosomal; 1.
DR   Pfam; PF00542; Ribosomal_L12; 1.
DR   ProDom; PD001326; Ribosomal_L12; 1.
DR   TIGRFAMs; TIGR00855; L12; 1.
KW   Ribosomal protein; Complete proteome.
SQ   SEQUENCE   124 AA;  12520 MW;  271C3CF71F80269E CRC64;

  Query Match              0.8%;  Score 7;  DB 1;  Length 124;
  Best Local Similarity   100.0%;  Pred. No. 45;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        354 GVSAAAV 360
              |||||||
Db         32 GVSAAAV 38


RESULT 37
H2AV_STRPU
ID   H2AV_STRPU     STANDARD;      PRT;   125 AA.
AC   P08991;
DT   01-NOV-1988 (Rel. 09, Created)
DT   01-NOV-1988 (Rel. 09, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Histone H2A variant (Fragment).
GN   H2A.F/Z.
OS   Strongylocentrotus purpuratus (Purple sea urchin).
OC   Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
OC   Echinoidea; Euechinoidea; Echinacea; Echinoida; Strongylocentrotidae;
OC   Strongylocentrotus.
OX   NCBI_TaxID=7668;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=87231039; PubMed=2438657;
RA   Ernst S.G., Miller H., Brenner C.A., Nocente-Mcgrath C., Francis S.,
RA   McIsaac R.;
RT   "Characterization of a cDNA clone coding for a sea urchin histone H2A
RT   variant related to the H2A.F/Z histone protein in vertebrates.";
RL   Nucleic Acids Res. 15:4629-4644(1987).
CC   -!- FUNCTION: VARIANT HISTONES H2A ARE SYNTHESIZED THROUGHOUT THE
CC       CELL CYCLE AND ARE VERY DIFFERENT FROM CLASSICAL S-PHASE REGULATED
CC       H2A. THE EXACT FUNCTION OF VARIANT HISTONES H2A IS NOT KNOWN.
CC   -!- SUBUNIT: The nucleosome is an octamer containing two molecules
CC       each of H2A, H2B, H3 and H4. The octamer wraps approximately 146
CC       BP of DNA.
CC   -!- SUBCELLULAR LOCATION: Nuclear.
CC   -!- SIMILARITY: Belongs to the histone H2A family.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; X05547; CAA29061.1; -.
DR   PIR; S07392; S07392.
DR   InterPro; IPR004822; Histone_core.
DR   InterPro; IPR002119; Histone_H2A.
DR   Pfam; PF00125; histone; 1.
DR   PRINTS; PR00620; HISTONEH2A.
DR   ProDom; PD000522; Histone_H2A; 1.
DR   SMART; SM00414; H2A; 1.
DR   PROSITE; PS00046; HISTONE_H2A; 1.
KW   Chromosomal protein; Nucleosome core; Nuclear protein; DNA-binding;
KW   Multigene family.
FT   NON_TER       1      1
SQ   SEQUENCE   125 AA;  13164 MW;  F65A17FDF263823F CRC64;

  Query Match              0.8%;  Score 7;  DB 1;  Length 125;
  Best Local Similarity   100.0%;  Pred. No. 45;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        139 ELDSLIK 145
              |||||||
Db         93 ELDSLIK 99


RESULT 38
H2AV_CHICK
ID   H2AV_CHICK     STANDARD;      PRT;   127 AA.
AC   P02272;
DT   21-JUL-1986 (Rel. 01, Created)
DT   21-JUL-1986 (Rel. 01, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Histone H2A variant.
GN   H2AF.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=83221493; PubMed=6574451;
RA   Harvey R.P., Whiting J.A., Coles L.S., Krieg P.A., Wells J.R.E.;
RT   "H2A.F: an extremely variant histone H2A sequence expressed in the
RT   chicken embryo.";
```

```
RL   Proc. Natl. Acad. Sci. U.S.A. 80:2819-2823(1983).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=89160327; PubMed=2493634;
RA   Dalton S., Robins A.J., Harvey R.P., Wells J.R.E.;
RT   "Transcription from the intron-containing chicken histone H2A.F gene
RT   is not S-phase regulated.";
RL   Nucleic Acids Res. 17:1745-1756(1989).
CC   -!- FUNCTION: VARIANT HISTONES H2A ARE SYNTHESIZED THROUGHOUT THE
CC       CELL CYCLE AND ARE VERY DIFFERENT FROM CLASSICAL S-PHASE REGULATED
CC       H2A. THE EXACT FUNCTION OF VARIANT HISTONES H2A IS NOT KNOWN.
CC   -!- SUBUNIT: The nucleosome is an octamer containing two molecules
CC       each of H2A, H2B, H3 and H4. The octamer wraps approximately 146
CC       BP of DNA.
CC   -!- SUBCELLULAR LOCATION: Nuclear.
CC   -!- DEVELOPMENTAL STAGE: EXPRESSED IN THE CHICKEN EMBRYO.
CC   -!- SIMILARITY: Belongs to the histone H2A family.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; V00414; CAA23705.1; -.
DR   EMBL; X13894; CAA32094.1; ALT_SEQ.
DR   PIR; S03282; HSCH2F.
DR   InterPro; IPR004822; Histone_core.
DR   InterPro; IPR002119; Histone_H2A.
DR   Pfam; PF00125; histone; 1.
DR   PRINTS; PR00620; HISTONEH2A.
DR   ProDom; PD000522; Histone_H2A; 1.
DR   SMART; SM00414; H2A; 1.
DR   PROSITE; PS00046; HISTONE_H2A; 1.
KW   Chromosomal protein; Nucleosome core; Nuclear protein; DNA-binding;
KW   Multigene family; Embryo.
FT   INIT_MET      0      0
SQ   SEQUENCE   127 AA;  13377 MW;  865AC88F6832F437 CRC64;

  Query Match             0.8%;  Score 7;  DB 1;  Length 127;
  Best Local Similarity   100.0%;  Pred. No. 45;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        139 ELDSLIK 145
              |||||||
Db         95 ELDSLIK 101

RESULT 39
H2AZ_HUMAN
ID   H2AZ_HUMAN     STANDARD;      PRT;   127 AA.
AC   P17317;
DT   01-AUG-1990 (Rel. 15, Created)
DT   01-AUG-1990 (Rel. 15, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Histone H2A.z (H2A/z).
GN   H2AFZ OR H2AZ.
OS   Homo sapiens (Human),
OS   Mus musculus (Mouse),
OS   Rattus norvegicus (Rat), and
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606, 10090, 10116, 9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   SPECIES=Human, Rat, and Bovine; TISSUE=Brain;
RX   MEDLINE=88143983; PubMed=3344202;
RA   Hatch C.L., Bonner W.M.;
RT   "Sequence of cDNAs for mammalian H2A.Z, an evolutionarily diverged
RT   but highly conserved basal histone H2A isoprotein species.";
RL   Nucleic Acids Res. 16:1113-1124(1988).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   SPECIES=Human;
RX   MEDLINE=90368704; PubMed=1697587;
RA   Hatch C.L., Bonner W.M.;
RT   "The human histone H2A.Z gene. Sequence and regulation.";
RL   J. Biol. Chem. 265:15211-15218(1990).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   SPECIES=Mouse; STRAIN=C57BL/6;
RA   Rocha D., Carrier A., Anderson E., Botcherby M., Guenet J.-L.,
RA   Jordan B.;
RL   Submitted (OCT-1996) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: VARIANT HISTONES H2A ARE SYNTHESIZED THROUGHOUT THE
CC       CELL CYCLE AND ARE VERY DIFFERENT FROM CLASSICAL S-PHASE REGULATED
CC       H2A. THE EXACT FUNCTION OF VARIANT HISTONES H2A IS NOT KNOWN.
CC   -!- SUBUNIT: The nucleosome is an octamer containing two molecules
CC       each of H2A, H2B, H3 and H4. The octamer wraps approximately 146
CC       BP of DNA.
CC   -!- SUBCELLULAR LOCATION: Nuclear.
CC   -!- SIMILARITY: Belongs to the histone H2A family.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; X52317; CAA36553.1; -.
DR   EMBL; M37583; AAA35984.1; -.
DR   EMBL; L10138; AAC61625.1; -.
DR   EMBL; X52316; CAA36552.1; -.
DR   EMBL; X52318; CAA36554.1; -.
DR   EMBL; M37584; AAA41329.1; -.
DR   EMBL; M37585; AAA30566.1; -.
DR   EMBL; U70494; AAB09578.1; -.
DR   PIR; A35881; A35881.
DR   PIR; S03642; S03642.
DR   PIR; S03644; S03644.
DR   PDB; 1F66; 27-NOV-00.
DR   Genew; HGNC:4741; H2AFZ.
DR   MIM; 142763; -.
DR   MGD; MGI:1888388; H2afz.
DR   GO; GO:0005718; C:nucleosome; NAS.
DR   GO; GO:0003677; F:DNA binding activity; NAS.
DR   GO; GO:0007001; P:chromosome organization and biogenesis (sen. . .; NAS.
DR   GO; GO:0006334; P:nucleosome assembly; NAS.
DR   InterPro; IPR004822; Histone_core.
DR   InterPro; IPR002119; Histone_H2A.
DR   Pfam; PF00125; histone; 1.
DR   PRINTS; PR00620; HISTONEH2A.
DR   ProDom; PD000522; Histone_H2A; 1.
DR   SMART; SM00414; H2A; 1.
DR   PROSITE; PS00046; HISTONE_H2A; 1.
KW   Chromosomal protein; Nucleosome core; Nuclear protein; DNA-binding;
KW   Multigene family; 3D-structure.
FT   INIT_MET      0      0
SQ   SEQUENCE   127 AA;  13421 MW;  794215381845F35A CRC64;

  Query Match             0.8%;  Score 7;  DB 1;  Length 127;
  Best Local Similarity   100.0%;  Pred. No. 45;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        139 ELDSLIK 145
              |||||||
Db         95 ELDSLIK 101

RESULT 40
RL20_BIFLO
```

```
ID   RL20_BIFLO     STANDARD;      PRT;   127 AA.
AC   Q8G4L1;
DT   15-SEP-2003 (Rel. 42, Created)
DT   15-SEP-2003 (Rel. 42, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   50S ribosomal protein L20.
GN   RPLT OR BL1367.
OS   Bifidobacterium longum.
OC   Bacteria; Actinobacteria; Actinobacteridae; Bifidobacteriales;
OC   Bifidobacteriaceae; Bifidobacterium.
OX   NCBI_TaxID=216816;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NCC 2705;
RX   MEDLINE=22294977; PubMed=12381787;
RA   Schell M.A., Karmirantzou M., Snel B., Vilanova D., Berger B.,
RA   Pessi G., Zwahlen M.-C., Desiere F., Bork P., Delley M.,
RA   Pridmore R.D., Arigoni F.;
RT   "The genome sequence of Bifidobacterium longum reflects its adaptation
RT   to the human gastrointestinal tract.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:14422-14427(2002).
CC   -!- FUNCTION: This protein binds directly to 23s ribosomal RNA and is
CC       necessary for the in vitro assembly process of the 50s ribosomal
CC       subunit. It is not involved in the protein synthesizing functions
CC       of that subunit (By similarity).
CC   -!- SIMILARITY: BELONGS TO THE L20P FAMILY OF RIBOSOMAL PROTEINS.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; AE014766; AAN25168.1; -.
DR   HAMAP; MF_00382; -; 1.
DR   InterPro; IPR005813; L20.
DR   InterPro; IPR005812; L20_bact_org.
DR   Pfam; PF00453; Ribosomal_L20; 1.
DR   PRINTS; PR00062; RIBOSOMALL20.
DR   ProDom; PD002389; L20; 1.
DR   TIGRFAMs; TIGR01032; rplT_bact; 1.
DR   PROSITE; PS00937; RIBOSOMAL_L20; 1.
KW   Ribosomal protein; rRNA-binding; Complete proteome.
SQ   SEQUENCE   127 AA;  14605 MW;  2EF7D5D28B7BDD99 CRC64;

  Query Match             0.8%;  Score 7;  DB 1;  Length 127;
  Best Local Similarity   100.0%;  Pred. No. 45;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        135 LAGIELD 141
              |||||||
Db         85 LAGIELD 91

Search completed: February 17, 2004, 10:20:08
Job time : 39 secs
```

```
OM protein - protein search, using sw model

Run on:         February 17, 2004, 10:17:00 ; Search time 45 Seconds
                                          (without alignments)
                                          5315.884 Million cell updates/sec

Title:          US-10-069-799-5
Perfect score:  927
Sequence:       1 MSNINVIKSNIQAGLNSTKS..........SSNALQPITQPTQGILAPSV 927

Scoring table:  OLIGO
                Gapop 60.0 , Gapext 60.0

Searched:       830525 seqs, 258052604 residues

Word size :      0

Total number of hits satisfying chosen parameters:      830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database :       SPTREMBL_23:*
                1:  sp_archea:*
                2:  sp_bacteria:*
                3:  sp_fungi:*
                4:  sp_human:*
                5:  sp_invertebrate:*
                6:  sp_mammal:*
                7:  sp_mhc:*
                8:  sp_organelle:*
                9:  sp_phage:*
                10:  sp_plant:*
                11:  sp_rodent:*
                12:  sp_virus:*
                13:  sp_vertebrate:*
                14:  sp_unclassified:*
                15:  sp_rvirus:*
                16:  sp_bacteriap:*
                17:  sp_archeap:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.
```

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	826	89.1	927	2	Q93GI2	Q93gi2 moraxella b
2	14	1.5	998	2	Q46716	Q46716 escherichia
3	14	1.5	998	2	O85101	O85101 escherichia
4	14	1.5	998	2	P71223	P71223 escherichia
5	14	1.5	998	2	Q47262	Q47262 escherichia
6	14	1.5	998	2	Q47461	Q47461 escherichia
7	14	1.5	998	2	Q9LC58	Q9lc58 escherichia
8	14	1.5	1024	2	Q8GA40	Q8ga40 escherichia
9	14	1.5	1024	2	Q8G9Z4	Q8g9z4 escherichia
10	14	1.5	1024	16	Q8FE01	Q8fe01 escherichia
11	14	1.5	1049	2	Q9RCG8	Q9rcg8 pasteurella
12	14	1.5	1052	2	Q93NP0	Q93np0 actinobacil
13	12	1.3	987	2	Q8KWZ9	Q8kwz9 actinobacil
14	12	1.3	987	2	Q8KWZ6	Q8kwz6 actinobacil
15	11	1.2	956	2	Q93NP1	Q93np1 actinobacil
16	11	1.2	1022	2	Q93NP2	Q93np2 actinobacil
17	10	1.1	593	5	Q8MR60	Q8mr60 drosophila
18	10	1.1	874	2	Q9ZFH0	Q9zfh0 azotobacter
19	10	1.1	889	2	Q8RMZ8	Q8rmz8 azotobacter
20	10	1.1	946	2	Q9EV24	Q9ev24 mannheimia
21	10	1.1	953	2	Q9EV31	Q9ev31 pasteurella
22	10	1.1	953	2	Q9ETX2	Q9etx2 mannheimia
23	10	1.1	953	2	Q9EV27	Q9ev27 pasteurella
24	10	1.1	953	2	Q9EV29	Q9ev29 pasteurella
25	10	1.1	953	2	Q9EV25	Q9ev25 mannheimia
26	10	1.1	953	2	Q9ETG5	Q9etg5 pasteurella
27	10	1.1	953	2	Q9EV23	Q9ev23 mannheimia
28	10	1.1	953	2	Q9EV33	Q9ev33 pasteurella
29	10	1.1	953	2	Q9EV34	Q9ev34 pasteurella
30	10	1.1	953	2	Q9EV32	Q9ev32 pasteurella
31	10	1.1	953	2	Q9EV30	Q9ev30 pasteurella
32	10	1.1	953	2	Q9EV28	Q9ev28 pasteurella
33	10	1.1	953	2	Q9EV26	Q9ev26 mannheimia
34	10	1.1	954	2	Q9EUE1	Q9eue1 pasteurella
35	10	1.1	955	2	Q9EUD4	Q9eud4 pasteurella
36	10	1.1	955	2	Q9EV22	Q9ev22 pasteurella
37	10	1.1	1403	2	Q44494	Q44494 azotobacter
38	10	1.1	1424	5	Q9VL42	Q9vl42 drosophila
39	10	1.1	4936	16	Q8YKJ3	Q8ykj3 anabaena sp
40	9	1.0	259	16	Q9RDA8	Q9rda8 streptomyce
41	9	1.0	567	2	Q9R9H2	Q9r9h2 pseudomonas
42	8	0.9	121	16	Q8PH96	Q8ph96 xanthomonas
43	8	0.9	161	16	Q8UHG5	Q8uhg5 agrobacteri
44	8	0.9	180	10	Q8LHL1	Q8lhl1 oryza sativ
45	8	0.9	190	16	Q8D3W2	Q8d3w2 vibrio vuln
46	8	0.9	208	2	Q51868	Q51868 pasteurella
47	8	0.9	219	2	O05199	O05199 rhizobium m
48	8	0.9	233	2	Q51865	Q51865 pasteurella
49	8	0.9	234	16	Q8ZAF8	Q8zaf8 yersinia pe
50	8	0.9	235	16	Q8PJA1	Q8pja1 xanthomonas
51	8	0.9	270	16	Q8UA74	Q8ua74 agrobacteri
52	8	0.9	310	16	Q8KEI9	Q8kei9 chlorobium
53	8	0.9	311	16	Q8Y0F0	Q8y0f0 ralstonia s
54	8	0.9	314	10	Q9SLP4	Q9slp4 tulipa bake
55	8	0.9	317	16	Q8U9U5	Q8u9u5 agrobacteri
56	8	0.9	321	16	Q8ZK75	Q8zk75 salmonella
57	8	0.9	321	16	Q8Z156	Q8z156 salmonella
58	8	0.9	322	2	Q44223	Q44223 anabaena sp
59	8	0.9	341	16	Q8G3J1	Q8g3j1 bifidobacte
60	8	0.9	341	16	Q8F4G0	Q8f4g0 leptospira
61	8	0.9	344	2	Q9X5A3	Q9x5a3 treponema d
62	8	0.9	353	16	Q92VW6	Q92vw6 rhizobium m
63	8	0.9	359	16	Q92KC8	Q92kc8 rhizobium m
64	8	0.9	363	16	Q92KB8	Q92kb8 rhizobium m
65	8	0.9	368	16	Q92FU0	Q92fu0 listeria in
66	8	0.9	368	16	Q8YAV0	Q8yav0 listeria mo
67	8	0.9	375	2	Q9WWF2	Q9wwf2 agrobacteri
68	8	0.9	380	16	Q8UKU5	Q8uku5 agrobacteri
69	8	0.9	384	10	Q9FV87	Q9fv87 petroselinu
70	8	0.9	384	10	Q9FV86	Q9fv86 petroselinu
71	8	0.9	394	5	Q17900	Q17900 caenorhabdi
72	8	0.9	396	16	Q92LP8	Q92lp8 rhizobium m
73	8	0.9	397	17	Q8ZVD7	Q8zvd7 pyrobaculum
74	8	0.9	399	5	Q17898	Q17898 caenorhabdi
75	8	0.9	400	11	Q99KX3	Q99kx3 mus musculu
76	8	0.9	401	10	Q9LGS6	Q9lgs6 oryza sativ
77	8	0.9	431	16	Q8XYF5	Q8xyf5 ralstonia s
78	8	0.9	443	16	Q8ZN34	Q8zn34 salmonella
79	8	0.9	443	16	Q8Z4M1	Q8z4m1 salmonella
80	8	0.9	447	17	Q9HPQ2	Q9hpq2 halobacteri
81	8	0.9	469	17	Q9HN79	Q9hn79 halobacteri
82	8	0.9	474	2	Q48230	Q48230 haemophilus
83	8	0.9	477	16	Q92NL7	Q92nl7 rhizobium m
84	8	0.9	477	17	Q9HPQ1	Q9hpq1 halobacteri
85	8	0.9	480	10	Q9FN53	Q9fn53 arabidopsis
86	8	0.9	505	16	Q9K7C5	Q9k7c5 bacillus ha
87	8	0.9	506	5	Q26707	Q26707 trypanosoma
88	8	0.9	510	5	Q8SYJ0	Q8syj0 drosophila
89	8	0.9	511	16	Q8ENV8	Q8env8 oceanobacil

90	8	0.9	530	10	Q8W041	Q8w041 arabidopsis
91	8	0.9	532	2	Q9ZFG8	Q9zfg8 azotobacter
92	8	0.9	537	10	Q9AYT3	Q9ayt3 nicotiana t
93	8	0.9	539	16	Q92XT9	Q92xt9 rhizobium m
94	8	0.9	553	2	Q44493	Q44493 azotobacter
95	8	0.9	559	17	Q9HRT5	Q9hrt5 halobacteri
96	8	0.9	589	16	Q8YR01	Q8yr01 anabaena sp
97	8	0.9	607	11	Q91ZY4	Q91zy4 mus musculu
98	8	0.9	643	2	O68085	O68085 rhodobacter
99	8	0.9	679	17	Q8TYL7	Q8tyl7 methanopyru
100	8	0.9	704	5	Q9VIE2	Q9vie2 drosophila
101	8	0.9	802	16	Q8CPJ0	Q8cpj0 staphylococ
102	8	0.9	825	2	Q8KWA8	Q8kwa8 ruegeria sp
103	8	0.9	850	2	Q937V4	Q937v4 bordetella
104	8	0.9	850	2	Q937V8	Q937v8 bordetella
105	8	0.9	850	2	Q937W0	Q937w0 bordetella
106	8	0.9	850	2	Q937V9	Q937v9 bordetella
107	8	0.9	850	2	Q932Y7	Q932y7 bordetella
108	8	0.9	850	2	Q937W1	Q937w1 bordetella
109	8	0.9	891	2	Q8RMZ9	Q8rmz9 azotobacter
110	8	0.9	993	16	Q8YYQ5	Q8yyq5 anabaena sp
111	8	0.9	997	2	Q44492	Q44492 azotobacter
112	8	0.9	997	2	Q44495	Q44495 azotobacter
113	8	0.9	1000	15	Q8JBT8	Q8jbt8 human immun
114	8	0.9	1003	15	Q8AK15	Q8ak15 human immun
115	8	0.9	1055	2	Q43892	Q43892 actinobacil
116	8	0.9	1112	16	Q92UV3	Q92uv3 rhizobium m
117	8	0.9	1168	2	Q9AGD2	Q9agd2 azotobacter
118	8	0.9	1208	16	Q9PFI9	Q9pfi9 xylella fas
119	8	0.9	1296	16	Q9P9W1	Q9p9w1 xylella fas
120	8	0.9	1417	16	Q8YTQ9	Q8ytq9 anabaena sp
121	8	0.9	1706	2	Q9L469	Q91469 bordetella
122	8	0.9	1736	10	O23025	O23025 arabidopsis
123	8	0.9	1741	16	P73817	P73817 synechocyst
124	8	0.9	1839	2	Q44496	Q44496 azotobacter
125	8	0.9	1990	16	Q8U7N7	Q8u7n7 agrobacteri
126	8	0.9	2361	5	Q8IEJ4	Q8iej4 plasmodium
127	8	0.9	2747	2	Q9L800	Q91800 aeromonas s
128	8	0.9	3083	16	Q8YL10	Q8yl10 anabaena sp
129	7	0.8	17	6	Q28459	Q28459 mirounga an
130	7	0.8	17	6	Q27943	Q27943 arctocephal
131	7	0.8	17	6	Q28246	Q28246 canis famil
132	7	0.8	17	6	Q28455	Q28455 leptonychot
133	7	0.8	17	6	Q28817	Q28817 phoca vitul
134	7	0.8	17	6	Q28445	Q28445 hydrurga le
135	7	0.8	31	6	Q28493	Q28493 mirounga le
136	7	0.8	63	5	Q9GSH2	Q9gsh2 bombyx mori
137	7	0.8	63	16	Q92JF1	Q92jf1 rickettsia
138	7	0.8	73	9	Q8SDW7	Q8sdw7 bacteriopha
139	7	0.8	73	9	Q9G032	Q9g032 bacteriopha
140	7	0.8	73	9	Q9MBS7	Q9mbs7 staphylococ
141	7	0.8	73	16	Q932B3	Q932b3 staphylococ
142	7	0.8	75	8	Q9G9R8	Q9g9r8 thunnus ton
143	7	0.8	75	8	Q9G101	Q9g101 thunnus obe
144	7	0.8	75	8	Q9G6D1	Q9g6d1 thunnus ala
145	7	0.8	75	8	Q9G6C5	Q9g6c5 thunnus alb
146	7	0.8	75	8	Q9G6C7	Q9g6c7 thunnus mac
147	7	0.8	75	8	Q9G6C9	Q9g6c9 thunnus thy
148	7	0.8	75	8	Q9G9S0	Q9g9s0 thunnus atl
149	7	0.8	75	8	Q9G6D3	Q9g6d3 thunnus thy
150	7	0.8	79	16	Q9A3P6	Q9a3p6 caulobacter
151	7	0.8	79	16	Q8U5J3	Q8u5j3 agrobacteri
152	7	0.8	81	16	Q9A4J2	Q9a4j2 caulobacter
153	7	0.8	85	16	Q9I0H5	Q9i0h5 pseudomonas
154	7	0.8	90	11	Q8R029	Q8r029 mus musculu
155	7	0.8	92	10	Q93WB4	Q93wb4 arabidopsis
156	7	0.8	98	10	Q8W233	Q8w233 zea mays (m
157	7	0.8	102	12	Q85128	Q85128 human parvo
158	7	0.8	102	16	Q8PL98	Q8pl98 xanthomonas
159	7	0.8	102	16	Q8P9H7	Q8p9h7 xanthomonas
160	7	0.8	104	12	Q85152	Q85152 human parvo
161	7	0.8	104	16	Q8Z1X8	Q8z1x8 salmonella
162	7	0.8	104	16	Q8ZJA1	Q8zja1 yersinia pe
163	7	0.8	104	16	Q8FD03	Q8fd03 escherichia
164	7	0.8	104	16	Q8EK58	Q8ek58 shewanella
165	7	0.8	105	2	Q9LBB1	Q9lbb1 helicobacte
166	7	0.8	105	2	Q9RAL9	Q9ral9 methylocyst
167	7	0.8	105	10	Q8H3B5	Q8h3b5 oryza sativ
168	7	0.8	105	16	Q9KNZ5	Q9knz5 vibrio chol
169	7	0.8	107	12	Q85139	Q85139 human parvo
170	7	0.8	107	16	Q98NQ9	Q98nq9 rhizobium l
171	7	0.8	108	12	Q85143	Q85143 human parvo
172	7	0.8	109	13	Q90WX7	Q90wx7 rana catesb
173	7	0.8	110	5	Q20637	Q20637 caenorhabdi
174	7	0.8	110	16	Q9KET3	Q9ket3 bacillus ha
175	7	0.8	111	2	O06120	O06120 methylocyst
176	7	0.8	112	10	Q944K9	Q944k9 arabidopsis
177	7	0.8	112	16	Q8YRE0	Q8yre0 anabaena sp
178	7	0.8	112	17	O27980	O27980 archaeoglob
179	7	0.8	116	2	Q9LB77	Q9lb77 helicobacte
180	7	0.8	117	2	Q9LBA3	Q9lba3 helicobacte
181	7	0.8	120	9	Q8W5Z8	Q8w5z8 bacteriopha
182	7	0.8	121	2	Q9LB62	Q9lb62 helicobacte
183	7	0.8	121	2	Q93PW9	Q93pw9 shigella so
184	7	0.8	123	6	O62695	O62695 oryctolagus
185	7	0.8	125	11	Q8CFT8	Q8cft8 mus musculu
186	7	0.8	126	2	Q9LB90	Q9lb90 helicobacte
187	7	0.8	126	17	Q9HRR7	Q9hrr7 halobacteri
188	7	0.8	127	16	Q8G4L1	Q8g4l1 bifidobacte
189	7	0.8	128	5	Q8MVI7	Q8mvi7 boltenia vi
190	7	0.8	130	2	Q9LB52	Q9lb52 helicobacte
191	7	0.8	130	16	Q8F0G1	Q8f0g1 leptospira
192	7	0.8	132	2	Q9LBB9	Q9lbb9 helicobacte
193	7	0.8	133	10	O04338	O04338 arabidopsis
194	7	0.8	133	10	Q39194	Q39194 arabidopsis
195	7	0.8	133	16	Q97E52	Q97e52 clostridium
196	7	0.8	134	12	Q85178	Q85178 human parvo
197	7	0.8	135	2	Q9LBB5	Q9lbb5 helicobacte
198	7	0.8	135	2	Q9LB67	Q9lb67 helicobacte
199	7	0.8	136	2	Q9K4W1	Q9k4w1 streptomyce
200	7	0.8	138	11	Q8BW15	Q8bw15 mus musculu
201	7	0.8	139	16	Q8ENQ5	Q8enq5 oceanobacil
202	7	0.8	140	2	Q9LB94	Q9lb94 helicobacte
203	7	0.8	140	2	Q9LB59	Q9lb59 helicobacte
204	7	0.8	140	2	Q9LB71	Q9lb71 helicobacte
205	7	0.8	140	12	Q9J891	Q9j891 spodoptera
206	7	0.8	142	4	Q8WYQ3	Q8wyq3 homo sapien
207	7	0.8	143	2	Q9LB57	Q9lb57 helicobacte
208	7	0.8	145	2	Q9LB99	Q9lb99 helicobacte
209	7	0.8	147	10	Q9C674	Q9c674 arabidopsis
210	7	0.8	147	16	Q53898	Q53898 streptomyce
211	7	0.8	149	5	Q19502	Q19502 caenorhabdi
212	7	0.8	151	2	Q9LB65	Q9lb65 helicobacte
213	7	0.8	151	4	Q9Y6H1	Q9y6h1 homo sapien
214	7	0.8	151	17	Q97Z82	Q97z82 sulfolobus
215	7	0.8	153	11	Q9D1L0	Q9d1l0 mus musculu
216	7	0.8	157	16	Q8PJZ0	Q8pjz0 xanthomonas
217	7	0.8	161	11	Q8K3S8	Q8k3s8 mus musculu
218	7	0.8	164	10	Q9FP57	Q9fp57 oryza sativ
219	7	0.8	167	4	Q8TEF2	Q8tef2 homo sapien
220	7	0.8	168	2	Q59609	Q59609 neisseria g
221	7	0.8	169	12	Q912Q3	Q912q3 human parvo
222	7	0.8	169	12	Q912C4	Q912c4 human eryth
223	7	0.8	169	12	Q912P1	Q912p1 human parvo
224	7	0.8	169	12	Q912Q1	Q912q1 human parvo
225	7	0.8	169	12	Q912Q7	Q912q7 human parvo
226	7	0.8	169	12	Q912N7	Q912n7 human parvo
227	7	0.8	169	12	Q912P9	Q912p9 human parvo
228	7	0.8	169	12	Q912C1	Q912c1 human eryth
229	7	0.8	169	12	Q912P6	Q912p6 human parvo
230	7	0.8	169	12	Q912N9	Q912n9 human parvo
231	7	0.8	169	12	Q912Q5	Q912q5 human parvo
232	7	0.8	169	12	Q912N2	Q912n2 human parvo
233	7	0.8	169	12	Q912D0	Q912d0 human eryth
234	7	0.8	169	12	Q912C6	Q912c6 human eryth
235	7	0.8	169	12	Q912N5	Q912n5 human parvo

236	7	0.8	169	12	Q912P4	Q912p4 human parvo
237	7	0.8	170	5	Q9NER7	Q9ner7 caenorhabdi
238	7	0.8	172	16	Q92V01	Q92v01 rhizobium m
239	7	0.8	175	2	Q8GQY7	Q8gqy7 thiobacillu
240	7	0.8	175	2	Q8GQY3	Q8gqy3 thiobacillu
241	7	0.8	175	16	Q8E6X0	Q8e6x0 streptococc
242	7	0.8	175	16	Q8E1F2	Q8e1f2 streptococc
243	7	0.8	176	12	Q85182	Q85182 human parvo
244	7	0.8	177	10	Q8W1H0	Q8w1h0 triticum ae
245	7	0.8	177	16	Q8DFN3	Q8dfn3 vibrio vuln
246	7	0.8	178	2	Q8GQY4	Q8gqy4 thiobacillu
247	7	0.8	178	16	Q9F3M4	Q9f3m4 streptomyce
248	7	0.8	179	12	Q85174	Q85174 human parvo
249	7	0.8	181	2	O68403	O68403 escherichia
250	7	0.8	181	2	O70070	O70070 escherichia
251	7	0.8	181	2	O68404	O68404 escherichia
252	7	0.8	181	8	Q9TLU6	Q9tlu6 cyanidium c
253	7	0.8	181	12	Q85120	Q85120 human parvo
254	7	0.8	182	12	Q85147	Q85147 human parvo
255	7	0.8	186	12	Q85163	Q85163 human parvo
256	7	0.8	186	12	Q85132	Q85132 human parvo
257	7	0.8	187	16	Q8XNW0	Q8xnw0 clostridium
258	7	0.8	187	16	Q8D5Y0	Q8d5y0 vibrio vuln
259	7	0.8	189	12	Q8JLD1	Q8jld1 ectromelia
260	7	0.8	189	16	Q8ZFF8	Q8zff8 yersinia pe
261	7	0.8	190	17	Q8ZX11	Q8zx11 pyrobaculum
262	7	0.8	192	16	Q8XKV0	Q8xkv0 clostridium
263	7	0.8	198	12	Q85193	Q85193 human parvo
264	7	0.8	199	10	Q8VXD6	Q8vxd6 nannochloro
265	7	0.8	199	16	P73231	P73231 synechocyst
266	7	0.8	202	5	Q8IJZ2	Q8ijz2 plasmodium
267	7	0.8	204	2	Q9Z3G4	Q9z3g4 synechococc
268	7	0.8	211	10	Q9LQB6	Q9lqb6 arabidopsis
269	7	0.8	211	12	Q8QTY8	Q8qty8 foot-and-mo
270	7	0.8	213	16	Q9CBJ1	Q9cbj1 mycobacteri
271	7	0.8	215	17	Q8ZZY1	Q8zzy1 pyrobaculum
272	7	0.8	216	16	Q9RUL4	Q9rul4 deinococcus
273	7	0.8	217	16	Q8U8S1	Q8u8s1 agrobacteri
274	7	0.8	218	16	Q9PF65	Q9pf65 xylella fas
275	7	0.8	219	12	Q9W9Q6	Q9w9q6 human parvo
276	7	0.8	219	12	Q9W9B9	Q9w9b9 human parvo
277	7	0.8	219	12	Q9WAS9	Q9was9 human parvo
278	7	0.8	219	16	Q9HWX3	Q9hwx3 pseudomonas
279	7	0.8	219	16	P96437	P96437 rhizobium m
280	7	0.8	219	16	O66733	O66733 aquifex aeo
281	7	0.8	220	17	O26974	O26974 methanobact
282	7	0.8	221	2	Q9ZBA7	Q9zba7 streptomyce
283	7	0.8	223	5	Q95TD5	Q95td5 drosophila
284	7	0.8	225	16	Q8PD09	Q8pd09 xanthomonas
285	7	0.8	227	9	Q8HA71	Q8ha71 bacteriopha
286	7	0.8	227	16	Q9HW21	Q9hw21 pseudomonas
287	7	0.8	228	16	Q8PGF3	Q8pgf3 xanthomonas
288	7	0.8	229	9	Q9ZX28	Q9zx28 mycobacteri
289	7	0.8	231	12	Q9WAU4	Q9wau4 human parvo
290	7	0.8	231	12	Q9WAT0	Q9wat0 human parvo
291	7	0.8	231	12	Q9W8V3	Q9w8v3 human parvo
292	7	0.8	231	12	Q9W9Z9	Q9w9z9 human parvo
293	7	0.8	231	12	Q9WAS8	Q9was8 human parvo
294	7	0.8	231	12	Q9WAU5	Q9wau5 human parvo
295	7	0.8	233	16	Q9EWM6	Q9ewm6 streptomyce
296	7	0.8	234	10	Q9FFH2	Q9ffh2 arabidopsis
297	7	0.8	237	2	O33630	O33630 saccharopol
298	7	0.8	238	2	O69260	O69260 borrelia re
299	7	0.8	241	16	Q9RZN1	Q9rzn1 deinococcus
300	7	0.8	242	2	Q9EX91	Q9ex91 petrotoga m
301	7	0.8	242	5	Q8T925	Q8t925 tetrahymena
302	7	0.8	242	16	Q8ZKK9	Q8zkk9 salmonella
303	7	0.8	243	3	Q96UE4	Q96ue4 neurospora
304	7	0.8	243	10	Q9CAX0	Q9cax0 arabidopsis
305	7	0.8	244	10	Q949S4	Q949s4 arabidopsis
306	7	0.8	244	10	Q8L8V6	Q8l8v6 arabidopsis
307	7	0.8	246	8	Q95FD3	Q95fd3 riccia flui
308	7	0.8	248	2	O87368	O87368 staphylococ
309	7	0.8	249	16	Q8U6D0	Q8u6d0 agrobacteri
310	7	0.8	249	16	Q9RKH4	Q9rkh4 streptomyce
311	7	0.8	250	12	Q9DPM8	Q9dpm8 human parvo
312	7	0.8	250	12	Q9DPM5	Q9dpm5 human parvo
313	7	0.8	250	12	Q9DPN5	Q9dpn5 human parvo
314	7	0.8	250	12	Q9DPM9	Q9dpm9 human parvo
315	7	0.8	250	12	Q9DPP1	Q9dpp1 human parvo
316	7	0.8	250	12	Q9DPN2	Q9dpn2 human parvo
317	7	0.8	250	12	Q9DPN3	Q9dpn3 human parvo
318	7	0.8	250	12	Q9DPN8	Q9dpn8 human parvo
319	7	0.8	250	12	Q9DPN1	Q9dpn1 human parvo
320	7	0.8	250	12	Q9DPN9	Q9dpn9 human parvo
321	7	0.8	250	12	Q9DPN4	Q9dpn4 human parvo
322	7	0.8	250	12	Q9DPN7	Q9dpn7 human parvo
323	7	0.8	250	12	Q9DPM4	Q9dpm4 human parvo
324	7	0.8	250	12	Q9DPP2	Q9dpp2 human parvo
325	7	0.8	250	12	Q9DPM6	Q9dpm6 human parvo
326	7	0.8	250	12	Q9DPN0	Q9dpn0 human parvo
327	7	0.8	250	12	Q9DPM7	Q9dpm7 human parvo
328	7	0.8	250	12	Q9DPN6	Q9dpn6 human parvo
329	7	0.8	250	12	Q9DPM3	Q9dpm3 human parvo
330	7	0.8	250	12	Q9DPP0	Q9dpp0 human parvo
331	7	0.8	251	16	Q8ZRM2	Q8zrm2 salmonella
332	7	0.8	251	16	Q8Z983	Q8z983 salmonella
333	7	0.8	253	16	Q8NVX1	Q8nvx1 staphylococ
334	7	0.8	254	16	Q8XEE8	Q8xee8 escherichia
335	7	0.8	256	16	Q8XP43	Q8xp43 clostridium
336	7	0.8	259	10	Q38730	Q38730 arum macula
337	7	0.8	260	2	Q8RQU3	Q8rqu3 thermus the
338	7	0.8	260	5	Q95WI2	Q95wi2 anopheles a
339	7	0.8	260	10	Q38731	Q38731 arum macula
340	7	0.8	261	8	Q94SD6	Q94sd6 dactylopten
341	7	0.8	261	8	Q94SE9	Q94se9 mastacembel
342	7	0.8	261	8	Q94T14	Q94t14 caelorinchu
343	7	0.8	261	8	Q94TG3	Q94tg3 ijimaia dof
344	7	0.8	261	8	Q94SB0	Q94sb0 pagrus majo
345	7	0.8	261	8	Q94SI8	Q94si8 antigonia c
346	7	0.8	261	8	Q94XJ1	Q94xj1 neoceratodu
347	7	0.8	261	8	Q94TB3	Q94tb3 neoscopelus
348	7	0.8	261	8	Q94SR6	Q94sr6 rondeletia
349	7	0.8	261	8	Q94TI9	Q94ti9 chauliodus
350	7	0.8	261	8	Q94T87	Q94t87 diaphus spl
351	7	0.8	261	8	Q94SQ3	Q94sq3 hoplostethu
352	7	0.8	261	8	Q9G6P4	Q9g6p4 polymixia j
353	7	0.8	261	8	Q8SF50	Q8sf50 ictalurus p
354	7	0.8	261	8	Q94TF0	Q94tf0 chlorophtha
355	7	0.8	261	8	Q9B603	Q9b603 plecoglossu
356	7	0.8	261	8	Q94SW8	Q94sw8 exocoetus v
357	7	0.8	261	8	Q94T48	Q94t48 zu cristatu
358	7	0.8	261	8	Q94SK1	Q94sk1 zenopsis ne
359	7	0.8	261	8	Q9B629	Q9b629 anguilla ja
360	7	0.8	261	8	Q94SS9	Q94ss9 danaceticht
361	7	0.8	261	8	Q94T38	Q94t38 polymixia l
362	7	0.8	261	8	Q94TC6	Q94tc6 saurida und
363	7	0.8	261	8	Q9MIA3	Q9mia3 typhlonecte
364	7	0.8	261	8	Q94SL4	Q94sl4 zeus faber
365	7	0.8	261	8	Q9B9U9	Q9b9u9 aulopus jap
366	7	0.8	261	8	Q94SY1	Q94sy1 cololabis s
367	7	0.8	261	8	Q94T61	Q94t61 trachipteru
368	7	0.8	261	8	Q94SM7	Q94sm7 myripristis
369	7	0.8	261	8	Q94TD8	Q94td8 harpadon mi
370	7	0.8	261	8	Q8HQP0	Q8hqp0 acipenser t
371	7	0.8	261	8	Q8HQL5	Q8hql5 amia calva
372	7	0.8	261	8	Q8HMU9	Q8hmu9 gymnothorax
373	7	0.8	261	8	Q8HMT7	Q8hmt7 erpetoichth
374	7	0.8	261	8	Q8HMS7	Q8hms7 polypterus
375	7	0.8	261	8	Q8HMR7	Q8hmr7 polypterus
376	7	0.8	261	8	Q8HMQ4	Q8hmq4 polyodon sp
377	7	0.8	261	8	Q8HMP1	Q8hmp1 scaphirhync
378	7	0.8	261	8	Q8HMH6	Q8hmh6 bassozetus
379	7	0.8	261	8	Q8HM86	Q8hm86 lota lota.
380	7	0.8	261	8	Q8HM73	Q8hm73 lophius lit
381	7	0.8	261	8	Q8HM60	Q8hm60 lophius ame

382	7	0.8	261	8	Q8HM21	Q8hm21 caulophryne
383	7	0.8	261	8	Q8HM08	Q8hm08 melanocetus
384	7	0.8	261	8	Q8HLZ5	Q8hlz5 melanotaeni
385	7	0.8	261	8	Q8HLW9	Q8hlw9 oryzias lat
386	7	0.8	261	8	Q8HLV7	Q8hlv7 gambusia af
387	7	0.8	261	8	Q8HLU4	Q8hlu4 cetostoma r
388	7	0.8	261	8	Q8HLT1	Q8hlt1 eutaeniopho
389	7	0.8	261	8	Q8HLQ5	Q8hlq5 diretmoides
390	7	0.8	261	8	Q8HLP2	Q8hlp2 diretmus ar
391	7	0.8	261	8	Q8HLM9	Q8hlm9 anomalops k
392	7	0.8	261	8	Q8HLL6	Q8hll6 monocentris
393	7	0.8	261	8	Q8HLK3	Q8hlk3 beryx decad
394	7	0.8	261	8	Q8HLJ0	Q8hlj0 ostichthys
395	7	0.8	261	8	Q8HLH7	Q8hlh7 sargocentro
396	7	0.8	261	8	Q8HLG4	Q8hlg4 parazen pac
397	7	0.8	261	8	Q8HLD8	Q8hld8 allocyttus
398	7	0.8	261	8	Q8HLC5	Q8hlc5 neocyttus r
399	7	0.8	261	8	Q8HKZ6	Q8hkz6 emmelichthy
400	7	0.8	261	8	Q8HKY3	Q8hky3 pterocaesio
401	7	0.8	261	8	Q8HKT1	Q8hkt1 salarias fa
402	7	0.8	261	8	Q8HKQ5	Q8hkq5 aspasma min
403	7	0.8	261	8	Q8HKP2	Q8hkp2 rhyacichthy
404	7	0.8	261	8	Q8HKM9	Q8hkm9 eleotris ac
405	7	0.8	261	8	Q8HKL6	Q8hkl6 sufflamen f
406	7	0.8	261	8	Q8HCV9	Q8hcv9 fugu rubrip
407	7	0.8	261	16	Q8EQZ4	Q8eqz4 oceanobacil
408	7	0.8	262	12	Q8B9H8	Q8b9h8 rachiplusia
409	7	0.8	262	16	O69841	O69841 streptomyce
410	7	0.8	264	13	Q9W7F6	Q9w7f6 polyodon sp
411	7	0.8	267	10	Q8RYD3	Q8ryd3 arabidopsis
412	7	0.8	269	10	Q9SWC6	Q9swc6 capsicum an
413	7	0.8	269	16	P95001	P95001 mycobacteri
414	7	0.8	270	10	Q9XGG7	Q9xgg7 pisum sativ
415	7	0.8	272	17	Q8U3F7	Q8u3f7 pyrococcus
416	7	0.8	274	2	Q9X5S8	Q9x5s8 streptomyce
417	7	0.8	274	2	Q9S3U2	Q9s3u2 zymomonas m
418	7	0.8	276	10	Q8W4T7	Q8w4t7 medicago tr
419	7	0.8	276	10	Q40750	Q40750 phaseolus a
420	7	0.8	278	16	Q9I3H1	Q9i3h1 pseudomonas
421	7	0.8	279	5	Q8MR71	Q8mr71 drosophila
422	7	0.8	279	16	Q9A873	Q9a873 caulobacter
423	7	0.8	280	16	Q8E460	Q8e460 streptococc
424	7	0.8	280	16	Q8DYK4	Q8dyk4 streptococc
425	7	0.8	281	16	Q8Y4S3	Q8y4s3 listeria mo
426	7	0.8	282	5	O18092	O18092 caenorhabdi
427	7	0.8	284	16	Q8XZQ4	Q8xzq4 ralstonia s
428	7	0.8	284	16	Q8DLF6	Q8dlf6 synechococc
429	7	0.8	287	2	Q8GPH2	Q8gph2 pantoea agg
430	7	0.8	287	16	Q8ELP0	Q8elp0 oceanobacil
431	7	0.8	290	2	Q52218	Q52218 klebsiella
432	7	0.8	290	16	Q8D4G3	Q8d4g3 vibrio vuln
433	7	0.8	291	16	Q8ZKC9	Q8zkc9 salmonella
434	7	0.8	291	16	Q8Z1M7	Q8z1m7 salmonella
435	7	0.8	291	16	Q9X1N5	Q9x1n5 thermotoga
436	7	0.8	293	13	Q8AXY4	Q8axy4 brachydanio
437	7	0.8	294	5	Q8IIB6	Q8iib6 plasmodium
438	7	0.8	294	16	Q8XF74	Q8xf74 salmonella
439	7	0.8	296	2	Q8RTR1	Q8rtr1 pseudomonas
440	7	0.8	296	16	Q8FRU9	Q8fru9 corynebacte
441	7	0.8	297	16	P74131	P74131 synechocyst
442	7	0.8	298	10	O23367	O23367 arabidopsis
443	7	0.8	299	13	Q9DGH3	Q9dgh3 brachydanio
444	7	0.8	299	16	Q55820	Q55820 synechocyst
445	7	0.8	300	2	Q8GC81	Q8gc81 enterobacte
446	7	0.8	300	4	Q8WW57	Q8ww57 homo sapien
447	7	0.8	304	11	Q8C7W6	Q8c7w6 mus musculu
448	7	0.8	304	16	Q92RB7	Q92rb7 rhizobium m
449	7	0.8	305	4	Q96A60	Q96a60 homo sapien
450	7	0.8	305	10	Q8L4J8	Q8l4j8 arabidopsis
451	7	0.8	307	16	Q8CR46	Q8cr46 staphylococ
452	7	0.8	308	16	Q8NML1	Q8nml1 corynebacte
453	7	0.8	309	6	Q8WNU0	Q8wnu0 macaca fasc
454	7	0.8	311	10	Q9ZVU4	Q9zvu4 arabidopsis
455	7	0.8	311	16	Q9RKA0	Q9rka0 streptomyce
456	7	0.8	311	17	Q9YDE9	Q9yde9 aeropyrum p
457	7	0.8	312	4	Q92977	Q92977 homo sapien
458	7	0.8	312	4	Q99620	Q99620 homo sapien
459	7	0.8	312	5	Q19261	Q19261 caenorhabdi
460	7	0.8	314	2	Q52240	Q52240 staphylococ
461	7	0.8	314	10	Q9M9A3	Q9m9a3 arabidopsis
462	7	0.8	314	16	Q8XNN5	Q8xnn5 clostridium
463	7	0.8	315	9	Q8SDT6	Q8sdt6 bacteriopha
464	7	0.8	315	16	Q931X0	Q931x0 staphylococ
465	7	0.8	315	16	Q8FLW0	Q8flw0 corynebacte
466	7	0.8	315	16	Q8ECM6	Q8ecm6 shewanella
467	7	0.8	316	10	Q941A9	Q941a9 arabidopsis
468	7	0.8	317	16	Q9KEH0	Q9keh0 bacillus ha
469	7	0.8	318	10	O64946	O64946 cucumis sat
470	7	0.8	319	16	Q9I2F8	Q9i2f8 pseudomonas
471	7	0.8	319	16	Q8YU40	Q8yu40 anabaena sp
472	7	0.8	320	10	Q9CAA2	Q9caa2 arabidopsis
473	7	0.8	320	10	Q94CK9	Q94ck9 arabidopsis
474	7	0.8	320	16	Q8YNL3	Q8ynl3 anabaena sp
475	7	0.8	321	16	Q8DLI7	Q8dli7 synechococc
476	7	0.8	324	16	Q9I0D3	Q9i0d3 pseudomonas
477	7	0.8	324	16	Q92PN4	Q92pn4 rhizobium m
478	7	0.8	325	5	Q8T896	Q8t896 ciona intes
479	7	0.8	325	17	Q8TY21	Q8ty21 methanopyru
480	7	0.8	325	17	Q8TVV2	Q8tvv2 methanopyru
481	7	0.8	326	3	Q02461	Q02461 saccharomyc
482	7	0.8	328	5	Q8SZR6	Q8szr6 drosophila
483	7	0.8	329	2	P72466	P72466 streptomyce
484	7	0.8	331	5	Q9VJ31	Q9vj31 drosophila
485	7	0.8	331	11	Q8R4T3	Q8r4t3 rattus norv
486	7	0.8	334	16	Q9RL53	Q9rl53 streptomyce
487	7	0.8	335	2	Q9L6C3	Q9l6c3 streptomyce
488	7	0.8	335	2	Q9RR28	Q9rr28 streptomyce
489	7	0.8	336	5	Q9VQL9	Q9vql9 drosophila
490	7	0.8	337	16	Q8P4X0	Q8p4x0 xanthomonas
491	7	0.8	338	16	Q9KQI1	Q9kqi1 vibrio chol
492	7	0.8	338	16	Q9A6I1	Q9a6i1 caulobacter
493	7	0.8	338	17	O58526	O58526 pyrococcus
494	7	0.8	339	16	Q987Z3	Q987z3 rhizobium l
495	7	0.8	339	16	Q8YGC6	Q8ygc6 brucella me
496	7	0.8	340	11	O35125	O35125 mus musculu
497	7	0.8	342	2	Q8KZB7	Q8kzb7 helicobacte
498	7	0.8	343	4	Q8N6K6	Q8n6k6 homo sapien
499	7	0.8	343	5	O18636	O18636 musca domes
500	7	0.8	345	16	Q98KQ2	Q98kq2 rhizobium l
501	7	0.8	346	8	Q9ML74	Q9ml74 bonasa sewe
502	7	0.8	346	8	Q9ML73	Q9ml73 bonasa umbe
503	7	0.8	346	16	Q9RKT4	Q9rkt4 streptomyce
504	7	0.8	348	16	Q9K9N0	Q9k9n0 bacillus ha
505	7	0.8	349	5	Q22512	Q22512 caenorhabdi
506	7	0.8	349	5	Q965U4	Q965u4 caenorhabdi
507	7	0.8	349	16	Q9FBZ8	Q9fbz8 streptomyce
508	7	0.8	349	16	Q8ZBG6	Q8zbg6 yersinia pe
509	7	0.8	349	16	Q8FSN3	Q8fsn3 corynebacte
510	7	0.8	350	16	Q8YNL0	Q8ynl0 anabaena sp
511	7	0.8	351	16	Q9KBZ3	Q9kbz3 bacillus ha
512	7	0.8	351	16	Q8ZDT2	Q8zdt2 yersinia pe
513	7	0.8	351	17	Q97Z26	Q97z26 sulfolobus
514	7	0.8	352	2	Q8KZC4	Q8kzc4 helicobacte
515	7	0.8	352	12	O73518	O73518 potato viru
516	7	0.8	352	17	Q9HKB6	Q9hkb6 thermoplasm
517	7	0.8	353	2	Q9S3L8	Q9s3l8 borrelia re
518	7	0.8	353	2	Q8KZA7	Q8kza7 helicobacte
519	7	0.8	355	8	O47483	O47483 cychrus oka
520	7	0.8	355	12	O73519	O73519 potato viru
521	7	0.8	355	12	O73510	O73510 potato viru
522	7	0.8	355	12	O73532	O73532 potato viru
523	7	0.8	355	12	O73533	O73533 potato viru
524	7	0.8	355	12	O73515	O73515 potato viru
525	7	0.8	355	12	O73520	O73520 potato viru
526	7	0.8	355	12	O73517	O73517 potato viru
527	7	0.8	357	5	O17591	O17591 caenorhabdi

528	7	0.8	358	16	O05579	O05579 mycobacteri
529	7	0.8	358	16	Q8Y1P7	Q8y1p7 ralstonia s
530	7	0.8	360	16	Q8YY68	Q8yy68 anabaena sp
531	7	0.8	362	2	Q8KZB5	Q8kzb5 helicobacte
532	7	0.8	363	2	Q9S3L9	Q9s3l9 borrelia re
533	7	0.8	364	16	Q8UIK8	Q8uik8 agrobacteri
534	7	0.8	364	16	Q8D549	Q8d549 vibrio vuln
535	7	0.8	365	10	Q8LDW7	Q8ldw7 arabidopsis
536	7	0.8	366	10	Q8S052	Q8s052 oryza sativ
537	7	0.8	367	9	O64040	O64040 bacteriopha
538	7	0.8	367	16	Q9KKV7	Q9kkv7 vibrio chol
539	7	0.8	367	16	O31982	O31982 bacillus su
540	7	0.8	368	16	Q9A048	Q9a048 streptococc
541	7	0.8	368	16	Q8K7T4	Q8k7t4 streptococc
542	7	0.8	370	2	Q9AE31	Q9ae31 rhizobium l
543	7	0.8	370	17	Q974H2	Q974h2 sulfolobus
544	7	0.8	371	2	Q9RLE3	Q9rle3 prochloroco
545	7	0.8	371	2	O86250	O86250 helicobacte
546	7	0.8	372	2	Q93K39	Q93k39 klebsiella
547	7	0.8	375	2	Q9AGC8	Q9agc8 enterobacte
548	7	0.8	375	2	Q9ALY0	Q9aly0 enterobacte
549	7	0.8	375	2	Q9AGC9	Q9agc9 enterobacte
550	7	0.8	375	2	Q9AGD1	Q9agd1 enterobacte
551	7	0.8	375	2	Q93T24	Q93t24 enterobacte
552	7	0.8	375	16	Q9RXS7	Q9rxs7 deinococcus
553	7	0.8	376	2	Q9AGD0	Q9agd0 enterobacte
554	7	0.8	376	2	O52496	O52496 bradyrhizob
555	7	0.8	376	17	Q96YZ5	Q96yz5 sulfolobus
556	7	0.8	377	2	Q8VRL8	Q8vrl8 thermus the
557	7	0.8	377	16	Q92SP9	Q92sp9 rhizobium m
558	7	0.8	379	16	Q9RV61	Q9rv61 deinococcus
559	7	0.8	379	16	Q8R790	Q8r790 thermoanaer
560	7	0.8	379	16	Q9RD86	Q9rd86 streptomyce
561	7	0.8	380	8	O48089	O48089 lacerta viv
562	7	0.8	380	16	Q9L0H9	Q9l0h9 streptomyce
563	7	0.8	381	10	Q9FJQ1	Q9fjq1 arabidopsis
564	7	0.8	381	16	Q8FUJ3	Q8fuj3 corynebacte
565	7	0.8	382	3	O93925	O93925 gibberella
566	7	0.8	382	16	O25891	O25891 helicobacte
567	7	0.8	384	3	O14094	O14094 schizosacch
568	7	0.8	384	10	Q9FV89	Q9fv89 petroselinu
569	7	0.8	384	10	Q9FV88	Q9fv88 petroselinu
570	7	0.8	385	12	Q8QLC1	Q8qlc1 mamestra co
571	7	0.8	385	16	Q92Z04	Q92z04 rhizobium m
572	7	0.8	386	16	O25780	O25780 helicobacte
573	7	0.8	386	16	Q9ZK50	Q9zk50 helicobacte
574	7	0.8	387	5	Q8IJS4	Q8ijs4 plasmodium
575	7	0.8	387	16	Q92VH1	Q92vh1 rhizobium m
576	7	0.8	388	16	Q8UJL8	Q8ujl8 agrobacteri
577	7	0.8	388	16	Q8UG48	Q8ug48 agrobacteri
578	7	0.8	389	2	Q9AIH9	Q9aih9 chlamydophi
579	7	0.8	389	16	Q8P970	Q8p970 xanthomonas
580	7	0.8	390	2	Q8KRN6	Q8krn6 cytophaga j
581	7	0.8	390	17	Q8TJL6	Q8tjl6 methanosarc
582	7	0.8	391	12	Q89469	Q89469 avian pneum
583	7	0.8	391	12	Q91HN5	Q91hn5 avian pneum
584	7	0.8	391	12	Q9DKS9	Q9dks9 avian pneum
585	7	0.8	391	16	Q8VKA2	Q8vka2 mycobacteri
586	7	0.8	392	5	Q9V6N7	Q9v6n7 drosophila
587	7	0.8	393	10	Q8S0N2	Q8s0n2 oryza sativ
588	7	0.8	393	16	Q8ZCI3	Q8zci3 yersinia pe
589	7	0.8	393	16	Q9EWE0	Q9ewe0 streptomyce
590	7	0.8	395	17	Q97VF3	Q97vf3 sulfolobus
591	7	0.8	396	2	Q93KV6	Q93kv6 streptomyce
592	7	0.8	396	2	Q93A40	Q93a40 enterococcu
593	7	0.8	397	16	Q9I223	Q9i223 pseudomonas
594	7	0.8	399	16	Q8DGM0	Q8dgm0 synechococc
595	7	0.8	400	16	Q9PHA8	Q9pha8 xylella fas
596	7	0.8	401	10	Q9LWM5	Q9lwm5 oryza sativ
597	7	0.8	401	16	Q9K7S7	Q9k7s7 bacillus ha
598	7	0.8	402	2	O33979	O33979 rhizobium e
599	7	0.8	403	2	O86253	O86253 helicobacte
600	7	0.8	404	17	Q97C00	Q97c00 thermoplasm

601	7	0.8	404	17	O29700	O29700 archaeoglob
602	7	0.8	406	5	Q8ILL1	Q8ill1 plasmodium
603	7	0.8	407	16	Q9CG50	Q9cg50 lactococcus
604	7	0.8	407	16	O51626	O51626 borrelia bu
605	7	0.8	408	2	O85454	O85454 pasteurella
606	7	0.8	408	16	Q9CMP3	Q9cmp3 pasteurella
607	7	0.8	408	16	Q9RCV5	Q9rcv5 streptomyce
608	7	0.8	408	16	Q8D0W6	Q8d0w6 yersinia pe
609	7	0.8	410	16	Q8X9A2	Q8x9a2 escherichia
610	7	0.8	411	16	Q92X82	Q92x82 rhizobium m
611	7	0.8	413	10	Q9FJS8	Q9fjs8 arabidopsis
612	7	0.8	414	13	Q98UJ2	Q98uj2 xenopus lae
613	7	0.8	414	16	O33415	O33415 pseudomonas
614	7	0.8	415	4	Q9UDU3	Q9udu3 homo sapien
615	7	0.8	415	16	Q8YUS7	Q8yus7 anabaena sp
616	7	0.8	418	11	Q63918	Q63918 mus musculu
617	7	0.8	418	11	Q8CBT4	Q8cbt4 mus musculu
618	7	0.8	422	16	Q8P4J2	Q8p4j2 xanthomonas
619	7	0.8	422	16	Q8CMQ6	Q8cmq6 staphylococ
620	7	0.8	424	5	O77237	O77237 drosophila
621	7	0.8	426	16	Q9A997	Q9a997 caulobacter
622	7	0.8	429	5	Q8I1Q5	Q8i1q5 plasmodium
623	7	0.8	429	15	P88173	P88173 human immun
624	7	0.8	429	16	Q97KC1	Q97kc1 clostridium
625	7	0.8	429	16	Q8Y0K6	Q8y0k6 ralstonia s
626	7	0.8	430	16	Q97FJ7	Q97fj7 clostridium
627	7	0.8	431	10	Q9M9S8	Q9m9s8 arabidopsis
628	7	0.8	431	16	Q8G4W4	Q8g4w4 bifidobacte
629	7	0.8	434	16	Q9L227	Q9l227 streptomyce
630	7	0.8	439	10	O65915	O65915 arabidopsis
631	7	0.8	439	16	Q8DKU1	Q8dku1 synechococc
632	7	0.8	440	16	Q8P988	Q8p988 xanthomonas
633	7	0.8	441	2	Q9X708	Q9x708 clostridium
634	7	0.8	441	16	Q9RSF5	Q9rsf5 deinococcus
635	7	0.8	442	10	Q8LME3	Q8lme3 oryza sativ
636	7	0.8	442	16	P96400	P96400 mycobacteri
637	7	0.8	443	11	Q8CFY5	Q8cfy5 mus musculu
638	7	0.8	443	16	Q8YYT6	Q8yyt6 anabaena sp
639	7	0.8	444	16	Q9HUR0	Q9hur0 pseudomonas
640	7	0.8	446	16	Q8ECB3	Q8ecb3 shewanella
641	7	0.8	449	2	Q9R6Z7	Q9r6z7 nocardioide
642	7	0.8	449	16	Q9CGJ0	Q9cgj0 lactococcus
643	7	0.8	449	16	Q9WYL2	Q9wyl2 thermotoga
644	7	0.8	449	16	Q8DFK6	Q8dfk6 vibrio vuln
645	7	0.8	450	10	Q9LPW3	Q9lpw3 arabidopsis
646	7	0.8	451	11	Q9D4L5	Q9d4l5 mus musculu
647	7	0.8	451	16	Q9K425	Q9k425 streptomyce
648	7	0.8	452	2	Q9X312	Q9x312 bacillus an
649	7	0.8	452	17	Q9V209	Q9v209 pyrococcus
650	7	0.8	452	17	Q97ZD7	Q97zd7 sulfolobus
651	7	0.8	454	16	Q8FAP1	Q8fap1 escherichia
652	7	0.8	454	17	O50131	O50131 pyrococcus
653	7	0.8	454	17	Q9V0Q7	Q9v0q7 pyrococcus
654	7	0.8	454	17	Q8U109	Q8u109 pyrococcus
655	7	0.8	455	16	Q9ZKC8	Q9zkc8 helicobacte
656	7	0.8	457	8	Q94QQ5	Q94qq5 caiman croc
657	7	0.8	458	16	O25134	O25134 helicobacte
658	7	0.8	460	4	Q9Y589	Q9y589 homo sapien
659	7	0.8	460	4	Q8TDF7	Q8tdf7 homo sapien
660	7	0.8	460	4	Q9NWZ3	Q9nwz3 homo sapien
661	7	0.8	460	11	Q62739	Q62739 rattus norv
662	7	0.8	460	16	Q9I5H2	Q9i5h2 pseudomonas
663	7	0.8	462	16	Q9KBI2	Q9kbi2 bacillus ha
664	7	0.8	463	16	Q8Z4Y1	Q8z4y1 salmonella
665	7	0.8	466	10	Q9LFH7	Q9lfh7 arabidopsis
666	7	0.8	469	5	Q8IJV7	Q8ijv7 plasmodium
667	7	0.8	471	16	Q9S2W9	Q9s2w9 streptomyce
668	7	0.8	472	2	Q8RNU1	Q8rnu1 helicobacte
669	7	0.8	473	2	Q9RB20	Q9rb20 pectobacter
670	7	0.8	473	16	Q8YCD4	Q8ycd4 brucella me
671	7	0.8	473	17	Q97WB8	Q97wb8 sulfolobus
672	7	0.8	474	16	Q8ZJS7	Q8zjs7 yersinia pe
673	7	0.8	476	17	Q8ZZR7	Q8zzr7 pyrobaculum

674	7	0.8	477	2	Q59388	Q59388 escherichia
675	7	0.8	477	16	Q9PBD6	Q9pbd6 xylella fas
676	7	0.8	478	2	Q9XB65	Q9xb65 erwinia amy
677	7	0.8	479	2	P95598	P95598 corynebacte
678	7	0.8	479	2	Q52080	Q52080 pseudomonas
679	7	0.8	480	4	Q8NDF9	Q8ndf9 homo sapien
680	7	0.8	480	16	Q8DLQ3	Q8dlq3 synechococc
681	7	0.8	482	5	Q9UA14	Q9ua14 dictyosteli
682	7	0.8	482	16	Q8UA11	Q8ua11 agrobacteri
683	7	0.8	482	17	O54610	O54610 halobacteri
684	7	0.8	483	5	Q9Y0A6	Q9y0a6 dictyosteli
685	7	0.8	483	5	Q8IJC0	Q8ijc0 plasmodium
686	7	0.8	483	16	Q9A8N5	Q9a8n5 caulobacter
687	7	0.8	484	16	Q9KRH2	Q9krh2 vibrio chol
688	7	0.8	485	5	O18125	O18125 caenorhabdi
689	7	0.8	488	17	Q57576	Q57576 methanococc
690	7	0.8	489	16	Q8XQL1	Q8xql1 ralstonia s
691	7	0.8	491	10	Q9FHE2	Q9fhe2 arabidopsis
692	7	0.8	491	17	Q8TMJ2	Q8tmj2 methanosarc
693	7	0.8	493	16	Q8XDS0	Q8xds0 escherichia
694	7	0.8	497	16	Q98I38	Q98i38 rhizobium l
695	7	0.8	498	5	Q9XU88	Q9xu88 caenorhabdi
696	7	0.8	498	16	Q8FVY2	Q8fvy2 brucella su
697	7	0.8	500	16	Q9KB49	Q9kb49 bacillus ha
698	7	0.8	500	16	Q8REU1	Q8reu1 fusobacteri
699	7	0.8	500	16	Q8F9S7	Q8f9s7 leptospira
700	7	0.8	500	17	Q978P9	Q978p9 thermoplasm
701	7	0.8	504	16	Q8DB29	Q8db29 vibrio vuln
702	7	0.8	505	10	O65558	O65558 arabidopsis
703	7	0.8	505	10	O48645	O48645 lycopersico
704	7	0.8	507	4	Q9P080	Q9p080 homo sapien
705	7	0.8	507	11	Q9Z125	Q9z125 mus musculu
706	7	0.8	507	12	Q9E1X3	Q9e1x3 cercopithec
707	7	0.8	509	2	Q9AN21	Q9an21 bradyrhizob
708	7	0.8	510	10	Q9LEX1	Q9lex1 arabidopsis
709	7	0.8	512	10	Q8LSK3	Q8lsk3 physcomitre
710	7	0.8	512	16	Q98N17	Q98n17 rhizobium l
711	7	0.8	512	16	P71828	P71828 mycobacteri
712	7	0.8	513	16	Q8FMH3	Q8fmh3 corynebacte
713	7	0.8	517	16	Q9RDI1	Q9rdi1 streptomyce
714	7	0.8	517	17	Q97BE6	Q97be6 thermoplasm
715	7	0.8	519	4	Q96BA8	Q96ba8 homo sapien
716	7	0.8	520	11	Q91W70	Q91w70 mus musculu
717	7	0.8	521	11	Q8CBN2	Q8cbn2 mus musculu
718	7	0.8	521	11	Q8CB61	Q8cb61 mus musculu
719	7	0.8	521	11	Q8BH52	Q8bh52 mus musculu
720	7	0.8	525	16	Q9A7F8	Q9a7f8 caulobacter
721	7	0.8	527	13	Q8AXY3	Q8axy3 fugu rubrip
722	7	0.8	527	13	Q8AXY2	Q8axy2 tetraodon n
723	7	0.8	529	16	Q8ET94	Q8et94 oceanobacil
724	7	0.8	530	5	O01983	O01983 caenorhabdi
725	7	0.8	531	12	Q9E3S5	Q9e3s5 human parvo
726	7	0.8	532	11	Q9R1W2	Q9r1w2 mus musculu
727	7	0.8	532	11	Q9WTL6	Q9wtl6 mus musculu
728	7	0.8	534	17	Q9YDN5	Q9ydn5 aeropyrum p
729	7	0.8	536	16	Q8FND6	Q8fnd6 corynebacte
730	7	0.8	537	5	Q8IPG6	Q8ipg6 drosophila
731	7	0.8	537	16	Q8XS42	Q8xs42 ralstonia s
732	7	0.8	537	16	Q9AK40	Q9ak40 streptomyce
733	7	0.8	538	16	Q9ADA6	Q9ada6 streptomyce
734	7	0.8	539	3	Q96U52	Q96u52 neurospora
735	7	0.8	543	5	O18556	O18556 musca domes
736	7	0.8	544	11	Q8BW24	Q8bw24 mus musculu
737	7	0.8	546	11	Q8BIL1	Q8bil1 mus musculu
738	7	0.8	548	16	Q92NZ9	Q92nz9 rhizobium m
739	7	0.8	548	16	Q9EWY9	Q9ewy9 streptomyce
740	7	0.8	550	16	Q98IS1	Q98is1 rhizobium l
741	7	0.8	551	5	Q9GYC4	Q9gyc4 leishmania
742	7	0.8	551	16	Q8NS49	Q8ns49 corynebacte
743	7	0.8	551	16	Q8FR55	Q8fr55 corynebacte
744	7	0.8	553	2	Q8KLJ4	Q8klj4 streptomyce
745	7	0.8	554	13	Q8AUN9	Q8aun9 fugu rubrip
746	7	0.8	557	16	Q8G1J5	Q8g1j5 brucella su
747	7	0.8	558	11	Q9WTL5	Q9wtl5 mus musculu
748	7	0.8	558	11	Q9R1W1	Q9r1w1 mus musculu
749	7	0.8	558	16	Q9X225	Q9x225 thermotoga
750	7	0.8	559	16	Q92VX5	Q92vx5 rhizobium m
751	7	0.8	561	5	Q9VZM5	Q9vzm5 drosophila
752	7	0.8	564	3	Q8J1V3	Q8j1v3 emericella
753	7	0.8	568	16	Q8DM73	Q8dm73 synechococc
754	7	0.8	575	10	Q8LGJ2	Q8lgj2 arabidopsis
755	7	0.8	575	10	Q9LZM8	Q9lzm8 arabidopsis
756	7	0.8	575	10	Q8H147	Q8h147 arabidopsis
757	7	0.8	575	16	Q8FW30	Q8fw30 brucella su
758	7	0.8	576	5	Q8ID23	Q8id23 plasmodium
759	7	0.8	576	16	P71664	P71664 mycobacteri
760	7	0.8	577	16	Q9L2F3	Q9l2f3 streptomyce
761	7	0.8	580	2	Q8KW93	Q8kw93 ruegeria sp
762	7	0.8	581	5	Q23969	Q23969 drosophila
763	7	0.8	581	10	Q8RX72	Q8rx72 arabidopsis
764	7	0.8	583	11	Q8CEW7	Q8cew7 mus musculu
765	7	0.8	584	16	Q8KCC2	Q8kcc2 chlorobium
766	7	0.8	587	5	Q8ITT6	Q8itt6 manduca sex
767	7	0.8	587	16	Q9KR30	Q9kr30 vibrio chol
768	7	0.8	588	2	Q8G9X8	Q8g9x8 escherichia
769	7	0.8	588	10	Q9MA86	Q9ma86 arabidopsis
770	7	0.8	589	5	Q94216	Q94216 caenorhabdi
771	7	0.8	589	11	Q8VDL5	Q8vdl5 mus musculu
772	7	0.8	589	16	Q92HC1	Q92hc1 rickettsia
773	7	0.8	590	10	Q8S4W7	Q8s4w7 vitis vinif
774	7	0.8	592	4	Q8NC78	Q8nc78 homo sapien
775	7	0.8	595	5	O45934	O45934 caenorhabdi
776	7	0.8	600	16	Q9KBX2	Q9kbx2 bacillus ha
777	7	0.8	605	3	Q96VA3	Q96va3 aspergillus
778	7	0.8	606	16	Q8U7N6	Q8u7n6 agrobacteri
779	7	0.8	608	16	Q9PR80	Q9pr80 ureaplasma
780	7	0.8	609	16	Q9PI93	Q9pi93 campylobact
781	7	0.8	610	2	Q9S0M8	Q9s0m8 deinococcus
782	7	0.8	613	5	Q8SWS5	Q8sws5 drosophila
783	7	0.8	616	16	Q8UAV6	Q8uav6 agrobacteri
784	7	0.8	617	2	Q9EVD8	Q9evd8 pseudomonas
785	7	0.8	617	2	Q9EVB2	Q9evb2 pseudomonas
786	7	0.8	617	16	Q9RUN0	Q9run0 deinococcus
787	7	0.8	620	16	Q9WYB6	Q9wyb6 thermotoga
788	7	0.8	621	11	Q8CAA7	Q8caa7 mus musculu
789	7	0.8	621	16	Q8EUH7	Q8euh7 mycoplasma
790	7	0.8	624	13	Q90WU2	Q90wu2 fundulus he
791	7	0.8	630	5	Q9VP21	Q9vp21 drosophila
792	7	0.8	632	16	Q928V3	Q928v3 listeria in
793	7	0.8	632	16	Q8Y4U6	Q8y4u6 listeria mo
794	7	0.8	634	5	Q9VRN6	Q9vrn6 drosophila
795	7	0.8	634	5	Q95PL8	Q95pl8 trypanosoma
796	7	0.8	634	17	Q8PRY4	Q8pry4 methanosarc
797	7	0.8	635	16	Q98QP3	Q98qp3 mycoplasma
798	7	0.8	636	4	Q8N7M8	Q8n7m8 homo sapien
799	7	0.8	641	4	Q9BXI0	Q9bxi0 homo sapien
800	7	0.8	641	7	Q8SNA1	Q8sna1 homo sapien
801	7	0.8	641	7	Q8MHQ1	Q8mhq1 homo sapien
802	7	0.8	642	17	Q8TIE9	Q8tie9 methanosarc
803	7	0.8	642	17	Q8PYV1	Q8pyv1 methanosarc
804	7	0.8	643	16	Q8G292	Q8g292 brucella su
805	7	0.8	646	5	Q9XU22	Q9xu22 caenorhabdi
806	7	0.8	652	16	Q98IC4	Q98ic4 rhizobium l
807	7	0.8	654	16	Q9CHI7	Q9chi7 lactococcus
808	7	0.8	655	16	O69855	O69855 streptomyce
809	7	0.8	658	2	Q8VP71	Q8vp71 klebsiella
810	7	0.8	658	10	Q9CAN3	Q9can3 arabidopsis
811	7	0.8	659	2	Q8GH79	Q8gh79 chlamydophi
812	7	0.8	659	16	Q8PPG6	Q8ppg6 xanthomonas
813	7	0.8	659	16	Q8P5D9	Q8p5d9 xanthomonas
814	7	0.8	662	5	O44832	O44832 caenorhabdi
815	7	0.8	670	16	Q8E0X1	Q8e0x1 streptococc
816	7	0.8	671	16	Q9KVH9	Q9kvh9 vibrio chol
817	7	0.8	674	4	Q9HA83	Q9ha83 homo sapien
818	7	0.8	674	11	Q8BZF7	Q8bzf7 mus musculu
819	7	0.8	676	16	Q8YVF5	Q8yvf5 anabaena sp

820	7	0.8	678	2	Q54679	Q54679 streptomyce
821	7	0.8	679	5	Q9U0V8	Q9u0v8 leishmania
822	7	0.8	687	2	Q9XDI8	Q9xdi8 bacteroides
823	7	0.8	689	16	Q927H1	Q927h1 listeria in
824	7	0.8	689	16	Q8Y410	Q8y410 listeria mo
825	7	0.8	691	4	Q8TC07	Q8tc07 homo sapien
826	7	0.8	692	17	Q8PVG2	Q8pvg2 methanosarc
827	7	0.8	693	5	Q21649	Q21649 caenorhabdi
828	7	0.8	695	16	Q8XWW4	Q8xww4 ralstonia s
829	7	0.8	701	4	Q96KS1	Q96ks1 homo sapien
830	7	0.8	708	10	Q8S555	Q8s555 picea abies
831	7	0.8	708	12	Q9DKW1	Q9dkw1 walrus cali
832	7	0.8	711	16	Q99S17	Q99s17 staphylococ
833	7	0.8	718	5	Q9BI07	Q9bi07 entamoeba h
834	7	0.8	722	16	Q9F322	Q9f322 streptomyce
835	7	0.8	726	16	Q8XRZ9	Q8xrz9 ralstonia s
836	7	0.8	726	17	Q9YE25	Q9ye25 aeropyrum p
837	7	0.8	735	16	P74453	P74453 synechocyst
838	7	0.8	739	10	Q9LWY3	Q9lwy3 oryza sativ
839	7	0.8	739	16	Q8FMZ8	Q8fmz8 corynebacte
840	7	0.8	745	16	Q8DHM6	Q8dhm6 synechococc
841	7	0.8	746	16	O66902	O66902 aquifex aeo
842	7	0.8	755	5	Q9VC61	Q9vc61 drosophila
843	7	0.8	756	16	Q9A888	Q9a888 caulobacter
844	7	0.8	757	16	Q9I4S8	Q9i4s8 pseudomonas
845	7	0.8	759	5	Q8IT91	Q8it91 ancylostoma
846	7	0.8	760	12	Q9PZT8	Q9pzt8 human parvo
847	7	0.8	761	12	Q9PZU0	Q9pzu0 human parvo
848	7	0.8	763	10	Q9SNP3	Q9snp3 oryza sativ
849	7	0.8	764	10	O49371	O49371 arabidopsis
850	7	0.8	765	3	Q9UVY1	Q9uvy1 pneumocysti
851	7	0.8	765	10	Q949H5	Q949h5 fragaria an
852	7	0.8	765	12	Q9PZT6	Q9pzt6 human parvo
853	7	0.8	767	16	Q92IB3	Q92ib3 rickettsia
854	7	0.8	773	12	Q913X1	Q913x1 human parvo
855	7	0.8	773	12	Q913W8	Q913w8 human parvo
856	7	0.8	781	12	Q8JN56	Q8jn56 human parvo
857	7	0.8	781	12	Q65789	Q65789 human parvo
858	7	0.8	781	12	P90223	P90223 human parvo
859	7	0.8	781	12	Q85191	Q85191 human parvo
860	7	0.8	781	12	P90221	P90221 human parvo
861	7	0.8	781	12	P89318	P89318 human parvo
862	7	0.8	781	12	Q8JYE3	Q8jye3 erythroviru
863	7	0.8	781	12	P90224	P90224 human parvo
864	7	0.8	781	12	Q9PZT0	Q9pzt0 human parvo
865	7	0.8	781	12	P89316	P89316 human parvo
866	7	0.8	781	12	P89320	P89320 human parvo
867	7	0.8	781	12	Q85117	Q85117 human parvo
868	7	0.8	781	12	P89321	P89321 human parvo
869	7	0.8	781	12	Q9JGS0	Q9jgs0 human parvo
870	7	0.8	781	12	P90222	P90222 human parvo
871	7	0.8	781	12	Q912B8	Q912b8 human eryth
872	7	0.8	781	12	Q9WKM0	Q9wkm0 human parvo
873	7	0.8	781	12	P89319	P89319 human parvo
874	7	0.8	781	12	Q90200	Q90200 human parvo
875	7	0.8	781	12	Q8JYD9	Q8jyd9 erythroviru
876	7	0.8	781	12	Q9JGP8	Q9jgp8 human parvo
877	7	0.8	781	12	P89317	P89317 human parvo
878	7	0.8	783	10	Q9C5L5	Q9c5l5 arabidopsis
879	7	0.8	783	10	Q8H1E5	Q8h1e5 arabidopsis
880	7	0.8	783	16	Q99QQ4	Q99qq4 streptomyce
881	7	0.8	790	16	Q8YSA9	Q8ysa9 anabaena sp
882	7	0.8	806	16	Q92YQ6	Q92yq6 rhizobium m
883	7	0.8	808	5	Q8T5K1	Q8t5k1 anopheles g
884	7	0.8	808	17	Q58796	Q58796 methanococc
885	7	0.8	815	16	P96644	P96644 bacillus su
886	7	0.8	816	15	Q994H2	Q994h2 human immun
887	7	0.8	818	5	Q9VS55	Q9vs55 drosophila
888	7	0.8	818	5	Q8TA44	Q8ta44 drosophila
889	7	0.8	822	4	Q96GV6	Q96gv6 homo sapien
890	7	0.8	824	11	Q8BXH3	Q8bxh3 mus musculu
891	7	0.8	826	2	Q9F9G4	Q9f9g4 helicobacte
892	7	0.8	829	2	Q9F9F8	Q9f9f8 helicobacte
893	7	0.8	829	2	Q9F9G0	Q9f9g0 helicobacte
894	7	0.8	833	5	Q9VLM1	Q9vlm1 drosophila
895	7	0.8	836	17	Q8U4M8	Q8u4m8 pyrococcus
896	7	0.8	839	2	Q93Q01	Q93q01 helicobacte
897	7	0.8	839	5	Q9NHN7	Q9nhn7 drosophila
898	7	0.8	839	5	Q9Y103	Q9y103 drosophila
899	7	0.8	843	5	Q9VY15	Q9vy15 drosophila
900	7	0.8	850	16	O66691	O66691 aquifex aeo
901	7	0.8	855	17	O57819	O57819 pyrococcus
902	7	0.8	855	17	Q9V2M1	Q9v2m1 pyrococcus
903	7	0.8	861	2	Q9F9G2	Q9f9g2 helicobacte
904	7	0.8	861	2	Q9F9F7	Q9f9f7 helicobacte
905	7	0.8	861	2	Q9F9G5	Q9f9g5 helicobacte
906	7	0.8	862	2	Q9F9G3	Q9f9g3 helicobacte
907	7	0.8	862	10	Q9FTQ0	Q9ftq0 oryza sativ
908	7	0.8	864	2	Q9F9G1	Q9f9g1 helicobacte
909	7	0.8	864	2	Q93Q00	Q93q00 helicobacte
910	7	0.8	866	2	Q9F9G6	Q9f9g6 helicobacte
911	7	0.8	866	5	Q9VF20	Q9vf20 drosophila
912	7	0.8	872	4	Q9H3R1	Q9h3r1 homo sapien
913	7	0.8	872	11	Q9EQW8	Q9eqw8 mus musculu
914	7	0.8	877	16	Q9RK68	Q9rk68 streptomyce
915	7	0.8	879	2	Q8KKY1	Q8kky1 rhizobium e
916	7	0.8	880	16	Q8FSG7	Q8fsg7 corynebacte
917	7	0.8	888	5	Q9VDD8	Q9vdd8 drosophila
918	7	0.8	893	4	Q9NZ12	Q9nz12 homo sapien
919	7	0.8	893	4	Q8TEA7	Q8tea7 homo sapien
920	7	0.8	896	5	Q8MQX4	Q8mqx4 drosophila
921	7	0.8	899	5	Q9NFX2	Q9nfx2 drosophila
922	7	0.8	899	5	Q9VGZ3	Q9vgz3 drosophila
923	7	0.8	899	5	O76935	O76935 drosophila
924	7	0.8	899	10	Q9C895	Q9c895 arabidopsis
925	7	0.8	900	3	O74925	O74925 schizosacch
926	7	0.8	900	16	Q8YZU3	Q8yzu3 anabaena sp
927	7	0.8	908	11	Q8C0J0	Q8c0j0 mus musculu
928	7	0.8	908	11	Q8BMI4	Q8bmi4 mus musculu
929	7	0.8	912	16	Q8P957	Q8p957 xanthomonas
930	7	0.8	928	11	Q8BW71	Q8bw71 mus musculu
931	7	0.8	929	5	Q9VR32	Q9vr32 drosophila
932	7	0.8	929	5	Q95XI7	Q95xi7 caenorhabdi
933	7	0.8	936	5	Q95VU2	Q95vu2 homarus ame
934	7	0.8	941	16	Q9ZIW5	Q9ziw5 streptomyce
935	7	0.8	942	11	Q8BG22	Q8bg22 mus musculu
936	7	0.8	946	5	Q9U6B5	Q9u6b5 dictyosteli
937	7	0.8	960	16	Q8Y377	Q8y377 ralstonia s
938	7	0.8	962	10	P93425	P93425 oryza sativ
939	7	0.8	977	5	Q9W3Y4	Q9w3y4 drosophila
940	7	0.8	982	5	Q9VUA5	Q9vua5 drosophila
941	7	0.8	991	11	Q8BX56	Q8bx56 mus musculu
942	7	0.8	993	2	Q8KUA3	Q8kua3 enterococcu
943	7	0.8	993	2	Q52054	Q52054 enterococcu
944	7	0.8	1000	5	O96696	O96696 heliothis v
945	7	0.8	1002	15	Q8JC77	Q8jc77 human immun
946	7	0.8	1005	15	Q8JBR1	Q8jbr1 human immun
947	7	0.8	1008	5	Q8MR76	Q8mr76 drosophila
948	7	0.8	1014	16	Q9JZG4	Q9jzg4 neisseria m
949	7	0.8	1014	16	Q9JUK9	Q9juk9 neisseria m
950	7	0.8	1016	4	Q92585	Q92585 homo sapien
951	7	0.8	1018	16	Q98LG7	Q98lg7 rhizobium l
952	7	0.8	1020	5	Q95XU2	Q95xu2 caenorhabdi
953	7	0.8	1022	15	Q90PW5	Q90pw5 chimpanzee
954	7	0.8	1024	5	Q8MRZ8	Q8mrz8 drosophila
955	7	0.8	1026	16	Q8EZW3	Q8ezw3 leptospira
956	7	0.8	1030	3	Q12180	Q12180 saccharomyc
957	7	0.8	1034	16	Q8FD25	Q8fd25 escherichia
958	7	0.8	1036	4	Q9NZV1	Q9nzv1 homo sapien
959	7	0.8	1038	13	Q90784	Q90784 gallus gall
960	7	0.8	1043	11	Q91Z79	Q91z79 rattus norv
961	7	0.8	1056	5	Q9W3H0	Q9w3h0 drosophila
962	7	0.8	1059	5	Q9U720	Q9u720 dictyosteli
963	7	0.8	1072	12	Q85672	Q85672 diadromus p
964	7	0.8	1072	16	Q92X83	Q92x83 rhizobium m
965	7	0.8	1077	13	Q8UWA6	Q8uwa6 tribolodon

966	7	0.8	1080	13	Q8AVN3	Q8avn3 xenopus lae
967	7	0.8	1082	2	Q9R964	Q9r964 helicobacte
968	7	0.8	1083	4	Q15386	Q15386 homo sapien
969	7	0.8	1083	11	Q8CDJ1	Q8cdj1 mus musculu
970	7	0.8	1083	11	Q8BQZ6	Q8bqz6 mus musculu
971	7	0.8	1085	5	Q8I5Z1	Q8i5z1 plasmodium
972	7	0.8	1088	13	Q8AYH7	Q8ayh7 oncorhynchu
973	7	0.8	1092	16	Q8EUS1	Q8eus1 mycoplasma
974	7	0.8	1100	12	Q91LA3	Q91la3 white spot
975	7	0.8	1109	2	Q53505	Q53505 campylobact
976	7	0.8	1109	2	Q8KYQ0	Q8kyq0 bacillus an
977	7	0.8	1112	2	O52781	O52781 campylobact
978	7	0.8	1117	10	Q9LT22	Q9lt22 arabidopsis
979	7	0.8	1122	16	Q8FW81	Q8fw81 brucella su
980	7	0.8	1131	16	Q9RV56	Q9rv56 deinococcus
981	7	0.8	1138	5	Q8IS12	Q8is12 dictyosteli
982	7	0.8	1140	17	Q9HK21	Q9hk21 thermoplasm
983	7	0.8	1146	3	Q06685	Q06685 saccharomyc
984	7	0.8	1149	17	Q8TPZ9	Q8tpz9 methanosarc
985	7	0.8	1155	10	Q8SB69	Q8sb69 oryza sativ
986	7	0.8	1155	12	Q9Q8W4	Q9q8w4 shope fibro
987	7	0.8	1155	12	Q9Q8I6	Q9q8i6 myxoma viru
988	7	0.8	1160	2	Q8RNY8	Q8rny8 pseudomonas
989	7	0.8	1163	2	Q45850	Q45850 clostridium
990	7	0.8	1165	16	Q98LX9	Q98lx9 rhizobium l
991	7	0.8	1170	16	Q8YC44	Q8yc44 brucella me
992	7	0.8	1174	5	Q8WRC4	Q8wrc4 tetrahymena
993	7	0.8	1180	17	Q8TI57	Q8ti57 methanosarc
994	7	0.8	1183	16	Q8PL49	Q8pl49 xanthomonas
995	7	0.8	1185	4	Q13135	Q13135 homo sapien
996	7	0.8	1191	10	Q9C873	Q9c873 arabidopsis
997	7	0.8	1202	4	Q8N4I2	Q8n4i2 homo sapien
998	7	0.8	1202	4	Q13136	Q13136 homo sapien
999	7	0.8	1204	5	Q93161	Q93161 caenorhabdi
1000	7	0.8	1208	12	O39273	O39273 equine herp

ALIGNMENTS

```
RESULT 1
Q93GI2
ID   Q93GI2      PRELIMINARY;     PRT;   927 AA.
AC   Q93GI2;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   RTX toxin.
GN   MBXA.
OS   Moraxella bovis.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC   Moraxellaceae; Moraxella.
OX   NCBI_TaxID=476;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Tifton I;
RX   MEDLINE=21388402; PubMed=11497442;
RA   Angelos J.A., Hess J.F., George L.W.;
RT   "Cloning and characterization of a Moraxella bovis cytotoxin gene.";
RL   Am. J. Vet. Res. 62:1222-1228(2001).
DR   EMBL; AF205359; AAK84651.1; -.
DR   InterPro; IPR002048; EF-hand.
DR   InterPro; IPR001343; Hemlysn_Ca_bind.
DR   InterPro; IPR003995; RtxA.
DR   Pfam; PF00353; hemolysinCabind; 5.
DR   Pfam; PF02382; RTX; 1.
DR   PRINTS; PR00313; CABNDNGRPT.
DR   PRINTS; PR01488; RTXTOXINA.
DR   PROSITE; PS00018; EF_HAND; 1.
DR   PROSITE; PS00330; HEMOLYSIN_CALCIUM; 1.
SQ   SEQUENCE   927 AA;  98845 MW;  F4B703577E10A96D CRC64;

  Query Match            89.1%;  Score 826;  DB 2;  Length 927;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches  926;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 MSNINVIKSNIQAGLNSTKSGLKNLYLAIPKDYDPQKGGTLNDFIKAADELGIARLAEEP 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MSNINVIKSNIQAGLNSTKSGLKNLYLAIPKDYDPQKGGTLNDFIKAADELGIARLAEEP 60

Qy         61 NHTETAKKSVDTVNQFLSLTQTGIAISATKLEKFLQKHSTNKLAKGLDSVENIDRKLGKA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 NHTETAKKSVDTVNQFLSLTQTGIAISATKLEKFLQKHSTNKLAKGLDSVENIDRKLGKA 120

Qy        121 SNVLSTLSSFLGTALAGIELDSLIKKGDAAPDALAKASIDLINEIIGNLSQSTQTIEAFS 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 SNVLSTLSSFLGTALAGIELDSLIKKGDAAPDALAKASIDLINEIIGNLSQSTQTIEAFS 180

Qy        181 SQLAKLGSTISQAKGFSNIGNKLQNLNFSKTNLGLEIITGLLSGISAGFALADKNASTGK 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 SQLAKLGSTISQAKGFSNIGNKLQNLNFSKTNLGLEIITGLLSGISAGFALADKNASTGK 240

Qy        241 KVAAGFELSNQVIGNVTKAISSYVLAQRVAAGLSTTGAVAALITSSIMLAISPLAFMNAA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 KVAAGFELSNQVIGNVTKAISSYVLAQRVAAGLSTTGAVAALITSSIMLAISPLAFMNAA 300

Qy        301 DKFNHANALDEFAKQFRKFGYDGDHLLAEYQRGVGTIEASLTTISTALGAVSAGVSAAAV 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 DKFNHANALDEFAKQFRKFGYDGDHLLAEYQRGVGTIEASLTTISTALGAVSAGVSAAAV 360

Qy        361 GSAVGTPIALLVAGVTGLISGILEASKQAMFESVANRLQGKILEWEKQNGGQNYFDKGYD 420
              ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GSAVGAPIALLVAGVTGLISGILEASKQAMFESVANRLQGKILEWEKQNGGQNYFDKGYD 420

Qy        421 SRYAAYLANNLKFLSELNKELEAERVIAITQQRWDNNIGELAGITKLGERIKSGKAYADA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 SRYAAYLANNLKFLSELNKELEAERVIAITQQRWDNNIGELAGITKLGERIKSGKAYADA 480

Qy        481 FEDGKKVEAGSNITLDAKTGIIDISNSNGKKTQALHFTSPLLTAGTESRERLTNGKYSYI 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 FEDGKKVEAGSNITLDAKTGIIDISNSNGKKTQALHFTSPLLTAGTESRERLTNGKYSYI 540

Qy        541 NKLKFGRVKNWQVTDGEASSKLDFSKVIQRVAETEGTDEIGLIVNAKAGNDDIFVGQGKM 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 NKLKFGRVKNWQVTDGEASSKLDFSKVIQRVAETEGTDEIGLIVNAKAGNDDIFVGQGKM 600

Qy        601 NIDGGDGHDRVFYSKDGGFGNITVDGTSATEAGSYTVNRKVARGDIYHEVVKRQETKVGK 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 NIDGGDGHDRVFYSKDGGFGNITVDGTSATEAGSYTVNRKVARGDIYHEVVKRQETKVGK 660

Qy        661 RTETIQYRDYELRKVGYGYQSTDNLKSVEEVIGSQFNDVFKGSKFNDIFHSGEGDDLLDG 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 RTETIQYRDYELRKVGYGYQSTDNLKSVEEVIGSQFNDVFKGSKFNDIFHSGEGDDLLDG 720

Qy        721 GAGDDRLFGGKGNDRLSGDEGDDLLDGGSGDDVLNGGAGNDVYIFRKGDGNDTLYDGTGN 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GAGDDRLFGGKGNDRLSGDEGDDLLDGGSGDDVLNGGAGNDVYIFRKGDGNDTLYDGTGN 780

Qy        781 DKLAFADANISDIMIERTKEGIIVKRNDHSGSINIPRWYITSNLQNYQSNKTDHKIEQLI 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 DKLAFADANISDIMIERTKEGIIVKRNDHSGSINIPRWYITSNLQNYQSNKTDHKIEQLI 840

Qy        841 GKDGSYITSDQIDKILQDKKDGTVITSQELKKLADENKSQKLSASDIASSLNKLVGSMAL 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GKDGSYITSDQIDKILQDKKDGTVITSQELKKLADENKSQKLSASDIASSLNKLVGSMAL 900

Qy        901 FGTANSVSSNALQPITQPTQGILAPSV 927
              |||||||||||||||||||||||||||
Db        901 FGTANSVSSNALQPITQPTQGILAPSV 927

RESULT 2
Q46716
ID   Q46716      PRELIMINARY;     PRT;   998 AA.
AC   Q46716;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
```

```
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Hemolysin A (HLYA).
GN   HLYA OR EHEC-HLYA.
OS   Escherichia coli O157:H7.
OG   Plasmid pO157.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Escherichia.
OX   NCBI_TaxID=83334;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98290540; PubMed=9628576;
RA   Makino K., Ishii K., Yasunaga T., Hattori M., Yokoyama K.,
RA   Yatsudo H.C., Kubota Y., Yamaichi Y., Iida T., Yamamoto K., Honda T.,
RA   Han C., Ohtsubo A., Kasamatsu M., Hayashi T., Kuhara S.,
RA   Shinagawa H.;
RT   "Complete nucleotide sequences of 93-kb and 3.3-kb plasmids of an
RT   enterohemorrhagic Escherichia coli O157:H7 derived from Sakai
RT   outbreak.";
RL   DNA Res. 5:1-9(1998).
RN   [2]
RP   SEQUENCE OF 241-998 FROM N.A.
RC   STRAIN=EHEC;
RA   Hall R.H., Xu J., Walderhaug M.O.;
RL   Submitted (JUL-1994) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=EDL933;
RX   MEDLINE=98391744; PubMed=9722640;
RA   Burland V., Shao Y., Perna N.T., Plunkett G., Sofia H.J.,
RA   Blattner F.R.;
RT   "The complete DNA sequence and analysis of the large virulence plasmid
RT   of Escherichia coli O157:H7.";
RL   Nucleic Acids Res. 26:4196-4204(1998).
DR   EMBL; AB011549; BAA31774.1; -.
DR   EMBL; U12572; AAA20544.1; -.
DR   EMBL; AF074613; AAC70116.1; -.
DR   InterPro; IPR001343; Hemlysn_Ca_bind.
DR   InterPro; IPR003995; RtxA.
DR   Pfam; PF00353; hemolysinCabind; 6.
DR   Pfam; PF02382; RTX; 1.
DR   PRINTS; PR00313; CABNDNGRPT.
DR   PRINTS; PR01488; RTXTOXINA.
KW   Plasmid.
SQ   SEQUENCE   998 AA;  107047 MW;  BB3C0D36FCBB9EBD CRC64;

  Query Match             1.5%;  Score 14;  DB 2;  Length 998;
  Best Local Similarity   100.0%;  Pred. No. 0.00015;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        379 ISGILEASKQAMFE 392
              ||||||||||||||
Db        392 ISGILEASKQAMFE 405

RESULT 3
O85101
ID   O85101      PRELIMINARY;      PRT;   998 AA.
AC   O85101;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Hemolysin.
GN   EHXA.
OS   Escherichia coli.
OG   Plasmid EHEC-hemolysin plasmid.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Escherichia.
OX   NCBI_TaxID=562;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=enterohemorrhagic EC920006;
RX   MEDLINE=98261495; PubMed=9596716;
RA   Boerlin P., Chen S., Colbourne J.K., Johnson R., De Grandis S.,
RA   Gyles C.;
RT   "Evolution of enterohemorrhagic escherichia coli hemolysin plasmids
RT   and the locus for enterocyte effacement in Shiga toxin-producing E.
RT   coli.";
RL   Infect. Immun. 66:2553-2561(1998).
DR   EMBL; AF043471; AAC24352.1; -.
DR   InterPro; IPR001343; Hemlysn_Ca_bind.
DR   InterPro; IPR003995; RtxA.
DR   Pfam; PF00353; hemolysinCabind; 6.
DR   Pfam; PF02382; RTX; 1.
DR   PRINTS; PR00313; CABNDNGRPT.
DR   PRINTS; PR01488; RTXTOXINA.
KW   Plasmid.
SQ   SEQUENCE   998 AA;  107095 MW;  E5A566B239DCBE55 CRC64;

  Query Match             1.5%;  Score 14;  DB 2;  Length 998;
  Best Local Similarity   100.0%;  Pred. No. 0.00015;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        379 ISGILEASKQAMFE 392
              ||||||||||||||
Db        392 ISGILEASKQAMFE 405

RESULT 4
P71223
ID   P71223      PRELIMINARY;      PRT;   998 AA.
AC   P71223;
DT   01-FEB-1997 (TrEMBLrel. 02, Created)
DT   01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   EHEC-hemolysin.
GN   EHEC-HLYA.
OS   Escherichia coli.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Escherichia.
OX   NCBI_TaxID=562;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=78-92;
RX   MEDLINE=97034824; PubMed=8880480;
RA   Schmidt H., Karch H.;
RT   "Enterohemolytic phenotypes and genoty[es pof Shiga Toxin-producing
RT   Esherichia coli O111 strain from patients with diarrhea and hemolytic-
RT   uremic syndrome.";
RL   J. Clin. Microbiol. 34:2364-2367(1996).
DR   EMBL; X94129; CAA63849.1; -.
DR   InterPro; IPR001343; Hemlysn_Ca_bind.
DR   InterPro; IPR003995; RtxA.
DR   Pfam; PF00353; hemolysinCabind; 6.
DR   Pfam; PF02382; RTX; 1.
DR   PRINTS; PR00313; CABNDNGRPT.
DR   PRINTS; PR01488; RTXTOXINA.
SQ   SEQUENCE   998 AA;  107255 MW;  DA3EF078C7E4131E CRC64;

  Query Match             1.5%;  Score 14;  DB 2;  Length 998;
  Best Local Similarity   100.0%;  Pred. No. 0.00015;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        379 ISGILEASKQAMFE 392
              ||||||||||||||
Db        392 ISGILEASKQAMFE 405

RESULT 5
Q47262
ID   Q47262      PRELIMINARY;      PRT;   998 AA.
AC   Q47262;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Hemolysin.
```

```
GN   EHEC-HLYA.
OS   Escherichia coli.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Escherichia.
OX   NCBI_TaxID=562;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=95172699; PubMed=7868227;
RA   Schmidt H., Beutin L., Karch H.;
RT   "Molecular analysis of the plasmid-encoded hemolysin of Escherichia
RT   coli 0157:H7 strain Edl933.";
RL   Infect. Immun. 63:1055-1061(1995).
DR   EMBL; X79839; CAA56234.1; -.
DR   InterPro; IPR001343; Hemlysn_Ca_bind.
DR   InterPro; IPR003995; RtxA.
DR   Pfam; PF00353; hemolysinCabind; 6.
DR   Pfam; PF02382; RTX; 1.
DR   PRINTS; PR00313; CABNDNGRPT.
DR   PRINTS; PR01488; RTXTOXINA.
SQ   SEQUENCE   998 AA;  107058 MW;  0D3BE108C309B8B3 CRC64;

  Query Match             1.5%;  Score 14;  DB 2;  Length 998;
  Best Local Similarity   100.0%;  Pred. No. 0.00015;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        379 ISGILEASKQAMFE 392
              ||||||||||||||
Db        392 ISGILEASKQAMFE 405

RESULT 6
Q47461
ID   Q47461      PRELIMINARY;      PRT;   998 AA.
AC   Q47461;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   EHEC-hlyA protein.
GN   EHEC-HLYA.
OS   Escherichia coli.
OG   Plasmid p0157.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Escherichia.
OX   NCBI_TaxID=562;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=EDL 933;
RX   MEDLINE=95172699; PubMed=7868227;
RA   Schmidt H., Beutin L., Karch H.;
RT   "Molecular analysis of the plasmid-encoded hemolysin of Escherichia
RT   coli 0157:H7 strain Edl933.";
RL   Infect. Immun. 63:1055-1061(1995).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=EDL 933;
RX   MEDLINE=97090409; PubMed=8936317;
RA   Schmidt H., Kernbach C., Karch H.;
RT   "Analysis of the EHEC hly operon and its location in the physical map
RT   of the large plasmid of enterohemorrhagic Esherichia coli 0157:H7.";
RL   Microbiology 142:907-914(1996).
DR   EMBL; X86087; CAA60042.1; -.
DR   InterPro; IPR001343; Hemlysn_Ca_bind.
DR   InterPro; IPR003995; RtxA.
DR   Pfam; PF00353; hemolysinCabind; 6.
DR   Pfam; PF02382; RTX; 1.
DR   PRINTS; PR00313; CABNDNGRPT.
DR   PRINTS; PR01488; RTXTOXINA.
KW   Plasmid.
SQ   SEQUENCE   998 AA;  107032 MW;  4D8BE108C309BF7E CRC64;

  Query Match             1.5%;  Score 14;  DB 2;  Length 998;
  Best Local Similarity   100.0%;  Pred. No. 0.00015;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        379 ISGILEASKQAMFE 392
              ||||||||||||||
Db        392 ISGILEASKQAMFE 405

RESULT 7
Q9LC58
ID   Q9LC58      PRELIMINARY;      PRT;   998 AA.
AC   Q9LC58;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Hemolysin A.
GN   EHEC-HLYA.
OS   Escherichia coli.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Escherichia.
OX   NCBI_TaxID=562;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Taneike I., Wakisaka-Saito N., Harada Y., Zhang H.-M., Yamamoto T.;
RT   "The enterohemorrhagic Escherichia coli (EHEC)-hemolysin genes of a
RT   Shiga toxin 1 (Stx1)- and Stx2-producing, serotype O128 Escherichia
RT   coli strain with a greatest hemolytic activity.";
RL   Acta Med. Biol. (Niigata) 0:0-0(2000).
DR   EMBL; AB032930; BAA93708.1; -.
DR   InterPro; IPR001343; Hemlysn_Ca_bind.
DR   InterPro; IPR003995; RtxA.
DR   Pfam; PF00353; hemolysinCabind; 6.
DR   Pfam; PF02382; RTX; 1.
DR   PRINTS; PR00313; CABNDNGRPT.
DR   PRINTS; PR01488; RTXTOXINA.
DR   PROSITE; PS00330; HEMOLYSIN_CALCIUM; 1.
SQ   SEQUENCE   998 AA;  107196 MW;  CD7A88E9BD862DB6 CRC64;

  Query Match             1.5%;  Score 14;  DB 2;  Length 998;
  Best Local Similarity   100.0%;  Pred. No. 0.00015;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        379 ISGILEASKQAMFE 392
              ||||||||||||||
Db        392 ISGILEASKQAMFE 405

RESULT 8
Q8GA40
ID   Q8GA40      PRELIMINARY;      PRT;  1024 AA.
AC   Q8GA40;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Hemolysin A.
GN   HLYA.
OS   Escherichia coli.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Escherichia.
OX   NCBI_TaxID=562;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=536;
RX   MEDLINE=22267134; PubMed=12379716;
RA   Dobrindt U., Blum-Oehler G., Nagy G., Schneider G., Johann A.,
RA   Gottschalk G., Hacker J.;
RT   "Genetic structure and distribution of four pathogenicity islands (PAI
RT   I536 to PAI IV536) of uropathogenic Escherichia coli strain 536.";
RL   Infect. Immun. 70:6365-6372(2002).
DR   EMBL; AJ488511; CAD33759.1; -.
SQ   SEQUENCE   1024 AA;  110314 MW;  9771F0A61ECF6EEB CRC64;

  Query Match             1.5%;  Score 14;  DB 2;  Length 1024;
  Best Local Similarity   100.0%;  Pred. No. 0.00015;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy        379 ISGILEASKQAMFE 392
              ||||||||||||||
Db        406 ISGILEASKQAMFE 419


RESULT 9
Q8G9Z4
ID   Q8G9Z4       PRELIMINARY;      PRT;  1024 AA.
AC   Q8G9Z4;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   HlyA protein.
GN   HLYA.
OS   Escherichia coli.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Escherichia.
OX   NCBI_TaxID=562;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=536;
RX   MEDLINE=22267134; PubMed=12379716;
RA   Dobrindt U., Blum-Oehler G., Nagy G., Schneider G., Johann A.,
RA   Gottschalk G., Hacker J.;
RT   "Genetic structure and distribution of four pathogenicity islands (PAI
RT   I536 to PAI IV536) of uropathogenic Escherichia coli strain 536.";
RL   Infect. Immun. 70:6365-6372(2002).
DR   EMBL; AJ494981; CAD42039.1; -.
SQ   SEQUENCE   1024 AA;  110365 MW;  3F3C284A45FEFEB3 CRC64;

  Query Match             1.5%;  Score 14;  DB 2;  Length 1024;
  Best Local Similarity   100.0%;  Pred. No. 0.00015;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        379 ISGILEASKQAMFE 392
              ||||||||||||||
Db        406 ISGILEASKQAMFE 419


RESULT 10
Q8FE01
ID   Q8FE01       PRELIMINARY;      PRT;  1024 AA.
AC   Q8FE01;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Hemolysin A.
GN   HLYA OR C3570.
OS   Escherichia coli O6.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Escherichia.
OX   NCBI_TaxID=217992;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=O6:H1 / CFT073 / ATCC 700928;
RX   MEDLINE=22388234; PubMed=12471157;
RA   Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,
RA   Rasko D., Buckles E.L., Liou S.-R., Boutin A., Hackett J., Stroud D.,
RA   Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,
RA   Mobley H.L.T., Donnenberg M.S., Blattner F.R.;
RT   "Extensive mosaic structure revealed by the complete genome sequence
RT   of uropathogenic Escherichia coli.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024(2002).
DR   EMBL; AE016766; AAN82018.1; -.
KW   Complete proteome.
SQ   SEQUENCE   1024 AA;  110246 MW;  9DEBA50C23B5A49C CRC64;

  Query Match             1.5%;  Score 14;  DB 16;  Length 1024;
  Best Local Similarity   100.0%;  Pred. No. 0.00015;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        379 ISGILEASKQAMFE 392
              ||||||||||||||
Db        406 ISGILEASKQAMFE 419


RESULT 11
Q9RCG8
ID   Q9RCG8       PRELIMINARY;      PRT;  1049 AA.
AC   Q9RCG8;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   PaxA.
GN   PAXA.
OS   Pasteurella aerogenes.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC   Pasteurellaceae; Pasteurella.
OX   NCBI_TaxID=749;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=field strain JF1319;
RA   Heyberger-Meyer B., Frey J., Nicolet J., Kuhnert P.;
RT   "Identification and characterization of a new RTX toxin operon (pax)
RT   in Pasteurella aerogenes.";
RL   Submitted (AUG-1996) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U66588; AAF15370.1; -.
DR   InterPro; IPR001343; Hemlysn_Ca_bind.
DR   InterPro; IPR003995; RtxA.
DR   Pfam; PF00353; hemolysinCabind; 6.
DR   Pfam; PF02382; RTX; 1.
DR   PRINTS; PR00313; CABNDNGRPT.
DR   PRINTS; PR01488; RTXTOXINA.
DR   PROSITE; PS00330; HEMOLYSIN_CALCIUM; 1.
SQ   SEQUENCE   1049 AA;  112308 MW;  9FA5070E48CC3127 CRC64;

  Query Match             1.5%;  Score 14;  DB 2;  Length 1049;
  Best Local Similarity   100.0%;  Pred. No. 0.00015;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        241 KVAAGFELSNQVIG 254
              ||||||||||||||
Db        275 KVAAGFELSNQVIG 288


RESULT 12
Q93NP0
ID   Q93NP0       PRELIMINARY;      PRT;  1052 AA.
AC   Q93NP0;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   RTX-toxin IIIA.
OS   Actinobacillus pleuropneumoniae (Haemophilus pleuropneumoniae).
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC   Pasteurellaceae; Actinobacillus.
OX   NCBI_TaxID=715;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Shin S.J., Park J.Y., Choi I.S., Shin N.R., Yoo H.S.;
RT   "Cloning and Sequencing of Apx IIIA from Actinobacillus
RT   pleuropneumoniae.";
RL   Submitted (MAR-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF363363; AAK50053.1; -.
DR   InterPro; IPR001343; Hemlysn_Ca_bind.
DR   InterPro; IPR003995; RtxA.
DR   Pfam; PF00353; hemolysinCabind; 6.
DR   Pfam; PF02382; RTX; 1.
DR   PRINTS; PR00313; CABNDNGRPT.
DR   PRINTS; PR01488; RTXTOXINA.
DR   PROSITE; PS00330; HEMOLYSIN_CALCIUM; 1.
SQ   SEQUENCE   1052 AA;  112862 MW;  CA0E160F02101F42 CRC64;

  Query Match             1.5%;  Score 14;  DB 2;  Length 1052;
  Best Local Similarity   100.0%;  Pred. No. 0.00015;
```

```
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        241 KVAAGFELSNQVIG 254
              ||||||||||||||
Db        275 KVAAGFELSNQVIG 288


RESULT 13
Q8KWZ9
ID   Q8KWZ9      PRELIMINARY;      PRT;   987 AA.
AC   Q8KWZ9;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   AqxA.
GN   AQXA.
OS   Actinobacillus cf. equuli.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC   Pasteurellaceae; Actinobacillus.
OX   NCBI_TaxID=182091;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CCUG 19799;
RX   MEDLINE=22031201; PubMed=12034544;
RA   Berthoud H., Frey J., Kuhnert P.;
RT   "Characterization of Aqx and its operon: the hemolytic RTX determinant
RT   of Actinobacillus equuli.";
RL   Vet. Microbiol. 87:159-174(2002).
DR   EMBL; AF381184; AAM45566.1; -.
DR   InterPro; IPR001343; Hemlysn_Ca_bind.
DR   InterPro; IPR003995; RtxA.
DR   Pfam; PF00353; hemolysinCabind; 1.
DR   Pfam; PF02382; RTX; 1.
DR   PRINTS; PR00313; CABNDNGRPT.
DR   PRINTS; PR01488; RTXTOXINA.
DR   PROSITE; PS00330; HEMOLYSIN_CALCIUM; 1.
SQ   SEQUENCE   987 AA;  105662 MW;  CC34CD62BF9B745D CRC64;

  Query Match             1.3%;  Score 12;  DB 2;  Length 987;
  Best Local Similarity   100.0%;  Pred. No. 0.016;
  Matches   12;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        243 AAGFELSNQVIG 254
              ||||||||||||
Db        277 AAGFELSNQVIG 288


RESULT 14
Q8KWZ6
ID   Q8KWZ6      PRELIMINARY;      PRT;   987 AA.
AC   Q8KWZ6;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   AqxA.
GN   AQXA.
OS   Actinobacillus equuli.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC   Pasteurellaceae; Actinobacillus.
OX   NCBI_TaxID=718;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=90-570;
RX   MEDLINE=22031201; PubMed=12034544;
RA   Berthoud H., Frey J., Kuhnert P.;
RT   "Characterization of Aqx and its operon: the hemolytic RTX determinant
RT   of Actinobacillus equuli.";
RL   Vet. Microbiol. 87:159-174(2002).
DR   EMBL; AF381185; AAM45569.1; -.
DR   InterPro; IPR001343; Hemlysn_Ca_bind.
DR   InterPro; IPR003995; RtxA.
DR   Pfam; PF00353; hemolysinCabind; 3.
DR   Pfam; PF02382; RTX; 1.
DR   PRINTS; PR00313; CABNDNGRPT.
DR   PRINTS; PR01488; RTXTOXINA.
SQ   SEQUENCE   987 AA;  105871 MW;  35440030AB522E53 CRC64;

  Query Match             1.3%;  Score 12;  DB 2;  Length 987;
  Best Local Similarity   100.0%;  Pred. No. 0.016;
  Matches   12;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        243 AAGFELSNQVIG 254
              ||||||||||||
Db        277 AAGFELSNQVIG 288


RESULT 15
Q93NP1
ID   Q93NP1      PRELIMINARY;      PRT;   956 AA.
AC   Q93NP1;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   RTX toxin IIA.
OS   Actinobacillus pleuropneumoniae (Haemophilus pleuropneumoniae).
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC   Pasteurellaceae; Actinobacillus.
OX   NCBI_TaxID=715;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Shin S.J., Park J.Y., Choi I.S., Shin N.R., Yoo H.S.;
RT   "Cloning and Sequencing of Apx IIA from Actinobacillus
RT   pleuropneumoniae.";
RL   Submitted (MAR-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF363362; AAK50052.1; -.
DR   InterPro; IPR001343; Hemlysn_Ca_bind.
DR   InterPro; IPR003995; RtxA.
DR   Pfam; PF00353; hemolysinCabind; 5.
DR   Pfam; PF02382; RTX; 1.
DR   PRINTS; PR00313; CABNDNGRPT.
DR   PRINTS; PR01488; RTXTOXINA.
DR   PROSITE; PS00330; HEMOLYSIN_CALCIUM; 1.
SQ   SEQUENCE   956 AA;  102495 MW;  9F8EB14EC76628D5 CRC64;

  Query Match             1.2%;  Score 11;  DB 2;  Length 956;
  Best Local Similarity   100.0%;  Pred. No. 0.17;
  Matches   11;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        369 ALLVAGVTGLI 379
              |||||||||||
Db        392 ALLVAGVTGLI 402


RESULT 16
Q93NP2
ID   Q93NP2      PRELIMINARY;      PRT;  1022 AA.
AC   Q93NP2;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   RTX toxin IA.
OS   Actinobacillus pleuropneumoniae (Haemophilus pleuropneumoniae).
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC   Pasteurellaceae; Actinobacillus.
OX   NCBI_TaxID=715;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Shin S.J., Park J.Y., Choi I.S., Shin N.R., Yoo H.S.;
RT   "Cloning and Sequencing of Apx IA from Actinobacillus
RT   pleuropneumoniae.";
RL   Submitted (MAR-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF363361; AAK50051.1; -.
DR   InterPro; IPR001343; Hemlysn_Ca_bind.
DR   InterPro; IPR003995; RtxA.
DR   Pfam; PF00353; hemolysinCabind; 6.
DR   Pfam; PF02382; RTX; 1.
```

```
DR   PRINTS; PR00313; CABNDNGRPT.
DR   PRINTS; PR01488; RTXTOXINA.
SQ   SEQUENCE   1022 AA;  110439 MW;  A68E486D335467E3 CRC64;

  Query Match             1.2%;  Score 11;  DB 2;  Length 1022;
  Best Local Similarity   100.0%;  Pred. No. 0.18;
  Matches   11;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        266 AQRVAAGLSTT 276
              |||||||||||
Db        289 AQRVAAGLSTT 299


RESULT 17
Q8MR60
ID   Q8MR60      PRELIMINARY;      PRT;   593 AA.
AC   Q8MR60;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   GH12004p.
GN   CG5833 OR CG13132.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Berkeley;
RA   Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA   Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
RA   George R., Gonzalez M., Guarin H., Kronmiller B., Li P., Liao G.,
RA   Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,
RA   Patel S., Phouanenavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA   Celniker S.;
RL   Submitted (JUN-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY122108; AAM52620.1; -.
DR   FlyBase; FBgn0032180; CG5833.
SQ   SEQUENCE   593 AA;  61037 MW;  5AA22B848C459627 CRC64;

  Query Match             1.1%;  Score 10;  DB 5;  Length 593;
  Best Local Similarity   100.0%;  Pred. No. 1.1;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        353 AGVSAAAVGS 362
              ||||||||||
Db        275 AGVSAAAVGS 284


RESULT 18
Q9ZFH0
ID   Q9ZFH0      PRELIMINARY;      PRT;   874 AA.
AC   Q9ZFH0;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Mannuronan C-5-epimerase AlgE6.
GN   ALGE6.
OS   Azotobacter vinelandii.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC   Pseudomonadaceae; Azotobacter.
OX   NCBI_TaxID=354;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=E;
RX   MEDLINE=99084941; PubMed=9864314;
RA   Svanem B.I.G., Skjaak-Braek G., Ertesvaag H., Valla S.;
RT   "Cloning and expression of three new Azotobacter vinelandii genes
RT   closely related to a previously described gene family encoding
RT   mannuronan C-5-epimerases.";
RL   J. Bacteriol. 181:68-77(1999).
DR   EMBL; AF099799; AAD04920.1; -.
DR   InterPro; IPR006633; CASH.
DR   InterPro; IPR001343; Hemlysn_Ca_bind.
DR   InterPro; IPR006626; PbH1.
DR   Pfam; PF00353; hemolysinCabind; 8.
DR   SMART; SM00722; CASH; 2.
DR   SMART; SM00710; PbH1; 8.
SQ   SEQUENCE   874 AA;  90159 MW;  22487F3A445427F1 CRC64;

  Query Match             1.1%;  Score 10;  DB 2;  Length 874;
  Best Local Similarity   100.0%;  Pred. No. 1.6;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        714 GDDLLDGGAG 723
              ||||||||||
Db        413 GDDLLDGGAG 422


RESULT 19
Q8RMZ8
ID   Q8RMZ8      PRELIMINARY;      PRT;   889 AA.
AC   Q8RMZ8;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Hypothetical 91.4 kDa protein.
OS   Azotobacter vinelandii.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC   Pseudomonadaceae; Azotobacter.
OX   NCBI_TaxID=354;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Svanem B.I., Valla S.;
RL   Submitted (FEB-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF486646; AAL93195.1; -.
DR   InterPro; IPR001343; Hemlysn_Ca_bind.
DR   Pfam; PF00353; hemolysinCabind; 8.
KW   Hypothetical protein.
SQ   SEQUENCE   889 AA;  91391 MW;  6945D3971E5410D5 CRC64;

  Query Match             1.1%;  Score 10;  DB 2;  Length 889;
  Best Local Similarity   100.0%;  Pred. No. 1.6;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        714 GDDLLDGGAG 723
              ||||||||||
Db        608 GDDLLDGGAG 617


RESULT 20
Q9EV24
ID   Q9EV24      PRELIMINARY;      PRT;   946 AA.
AC   Q9EV24;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Leukotoxin.
GN   LKTA.
OS   Mannheimia glucosida.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC   Pasteurellaceae; Mannheimia.
OX   NCBI_TaxID=85401;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PH574;
RX   MEDLINE=21101823; PubMed=11157953;
RA   Davies R.L., Whittam T.S., Selander R.K.;
RT   "Sequence Diversity and Molecular Evolution of the Leukotoxin (lktA)
RT   Gene in Bovine and Ovine Strains of Mannheimia (Pasteurella)
RT   haemolytica.";
RL   J. Bacteriol. 183:1394-1404(2001).
DR   EMBL; AF314521; AAG40305.1; -.
DR   InterPro; IPR001753; EnCoA_hydrtse.
DR   InterPro; IPR001343; Hemlysn_Ca_bind.
```

```
DR   InterPro; IPR003995; RtxA.
DR   Pfam; PF00353; hemolysinCabind; 5.
DR   Pfam; PF02382; RTX; 1.
DR   PRINTS; PR00313; CABNDNGRPT.
DR   PRINTS; PR01488; RTXTOXINA.
DR   PROSITE; PS00166; ENOYL_COA_HYDRATASE; 1.
SQ   SEQUENCE   946 AA;  101480 MW;  25C077858BDC76C4 CRC64;

  Query Match             1.1%;  Score 10;  DB 2;  Length 946;
  Best Local Similarity   100.0%;  Pred. No. 1.7;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        265 LAQRVAAGLS 274
              ||||||||||
Db        276 LAQRVAAGLS 285


RESULT 21
Q9EV31
ID   Q9EV31      PRELIMINARY;      PRT;   953 AA.
AC   Q9EV31;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Leukotoxin.
GN   LKTA.
OS   Pasteurella haemolytica.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC   Pasteurellaceae; Mannheimia.
OX   NCBI_TaxID=75985;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PH588;
RX   MEDLINE=21101823; PubMed=11157953;
RA   Davies R.L., Whittam T.S., Selander R.K.;
RT   "Sequence Diversity and Molecular Evolution of the Leukotoxin (lktA)
RT   Gene in Bovine and Ovine Strains of Mannheimia (Pasteurella)
RT   haemolytica.";
RL   J. Bacteriol. 183:1394-1404(2001).
DR   EMBL; AF314510; AAG40294.1; -.
DR   InterPro; IPR001753; EnCoA_hydrtse.
DR   InterPro; IPR001343; Hemlysn_Ca_bind.
DR   InterPro; IPR003995; RtxA.
DR   Pfam; PF00353; hemolysinCabind; 5.
DR   Pfam; PF02382; RTX; 1.
DR   PRINTS; PR00313; CABNDNGRPT.
DR   PRINTS; PR01488; RTXTOXINA.
DR   PROSITE; PS00166; ENOYL_COA_HYDRATASE; 1.
SQ   SEQUENCE   953 AA;  102160 MW;  A189BF80754A7907 CRC64;

  Query Match             1.1%;  Score 10;  DB 2;  Length 953;
  Best Local Similarity   100.0%;  Pred. No. 1.7;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        265 LAQRVAAGLS 274
              ||||||||||
Db        283 LAQRVAAGLS 292


RESULT 22
Q9ETX2
ID   Q9ETX2      PRELIMINARY;      PRT;   953 AA.
AC   Q9ETX2;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Leukotoxin.
GN   LKTA.
OS   Mannheimia glucosida.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC   Pasteurellaceae; Mannheimia.
OX   NCBI_TaxID=85401;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PH498, and PH344;
RA   Davies R.L., Whittam T.S., Selander R.K.;
RT   "Sequence diversity and molecular evolution of the leukotoxin (lktA)
RT   gene in bovine and ovine strains of Mannheimia (Pasteurella)
RT   haemolytica.";
RL   J. Bacteriol. 0:0-0(2001).
DR   EMBL; AF314518; AAG40302.1; -.
DR   EMBL; AF314517; AAG40301.1; -.
DR   InterPro; IPR001753; EnCoA_hydrtse.
DR   InterPro; IPR001343; Hemlysn_Ca_bind.
DR   InterPro; IPR003995; RtxA.
DR   Pfam; PF00353; hemolysinCabind; 5.
DR   Pfam; PF02382; RTX; 1.
DR   PRINTS; PR00313; CABNDNGRPT.
DR   PRINTS; PR01488; RTXTOXINA.
DR   PROSITE; PS00166; ENOYL_COA_HYDRATASE; 1.
SQ   SEQUENCE   953 AA;  102135 MW;  70DB354157F5881E CRC64;

  Query Match             1.1%;  Score 10;  DB 2;  Length 953;
  Best Local Similarity   100.0%;  Pred. No. 1.7;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        265 LAQRVAAGLS 274
              ||||||||||
Db        283 LAQRVAAGLS 292


RESULT 23
Q9EV27
ID   Q9EV27      PRELIMINARY;      PRT;   953 AA.
AC   Q9EV27;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Leukotoxin.
GN   LKTA.
OS   Pasteurella haemolytica.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC   Pasteurellaceae; Mannheimia.
OX   NCBI_TaxID=75985;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PH392;
RX   MEDLINE=21101823; PubMed=11157953;
RA   Davies R.L., Whittam T.S., Selander R.K.;
RT   "Sequence Diversity and Molecular Evolution of the Leukotoxin (lktA)
RT   Gene in Bovine and Ovine Strains of Mannheimia (Pasteurella)
RT   haemolytica.";
RL   J. Bacteriol. 183:1394-1404(2001).
DR   EMBL; AF314516; AAG40300.1; -.
DR   InterPro; IPR001753; EnCoA_hydrtse.
DR   InterPro; IPR001343; Hemlysn_Ca_bind.
DR   InterPro; IPR003995; RtxA.
DR   Pfam; PF00353; hemolysinCabind; 5.
DR   Pfam; PF02382; RTX; 1.
DR   PRINTS; PR00313; CABNDNGRPT.
DR   PRINTS; PR01488; RTXTOXINA.
DR   PROSITE; PS00166; ENOYL_COA_HYDRATASE; 1.
SQ   SEQUENCE   953 AA;  102230 MW;  2B686808EB370090 CRC64;

  Query Match             1.1%;  Score 10;  DB 2;  Length 953;
  Best Local Similarity   100.0%;  Pred. No. 1.7;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        265 LAQRVAAGLS 274
              ||||||||||
Db        283 LAQRVAAGLS 292


RESULT 24
Q9EV29
ID   Q9EV29      PRELIMINARY;      PRT;   953 AA.
```

```
AC   Q9EV29;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Leukotoxin.
GN   LKTA.
OS   Pasteurella haemolytica.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC   Pasteurellaceae; Mannheimia.
OX   NCBI_TaxID=75985;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PH278;
RX   MEDLINE=21101823; PubMed=11157953;
RA   Davies R.L., Whittam T.S., Selander R.K.;
RT   "Sequence Diversity and Molecular Evolution of the Leukotoxin (lktA)
RT   Gene in Bovine and Ovine Strains of Mannheimia (Pasteurella)
RT   haemolytica.";
RL   J. Bacteriol. 183:1394-1404(2001).
DR   EMBL; AF314514; AAG40298.1; -.
DR   InterPro; IPR001753; EnCoA_hydrtse.
DR   InterPro; IPR001343; Hemlysn_Ca_bind.
DR   InterPro; IPR003995; RtxA.
DR   Pfam; PF00353; hemolysinCabind; 5.
DR   Pfam; PF02382; RTX; 1.
DR   PRINTS; PR00313; CABNDNGRPT.
DR   PRINTS; PR01488; RTXTOXINA.
DR   PROSITE; PS00166; ENOYL_COA_HYDRATASE; 1.
SQ   SEQUENCE   953 AA;  102147 MW;  11600FDA7849A1CA CRC64;

  Query Match             1.1%;  Score 10;  DB 2;  Length 953;
  Best Local Similarity   100.0%;  Pred. No. 1.7;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        265 LAQRVAAGLS 274
              ||||||||||
Db        283 LAQRVAAGLS 292

RESULT 25
Q9EV25
ID   Q9EV25      PRELIMINARY;      PRT;   953 AA.
AC   Q9EV25;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Leukotoxin.
GN   LKTA.
OS   Mannheimia glucosida.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC   Pasteurellaceae; Mannheimia.
OX   NCBI_TaxID=85401;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PH496;
RX   MEDLINE=21101823; PubMed=11157953;
RA   Davies R.L., Whittam T.S., Selander R.K.;
RT   "Sequence Diversity and Molecular Evolution of the Leukotoxin (lktA)
RT   Gene in Bovine and Ovine Strains of Mannheimia (Pasteurella)
RT   haemolytica.";
RL   J. Bacteriol. 183:1394-1404(2001).
DR   EMBL; AF314520; AAG40304.1; -.
DR   InterPro; IPR001753; EnCoA_hydrtse.
DR   InterPro; IPR001343; Hemlysn_Ca_bind.
DR   InterPro; IPR003995; RtxA.
DR   Pfam; PF00353; hemolysinCabind; 5.
DR   Pfam; PF02382; RTX; 1.
DR   PRINTS; PR00313; CABNDNGRPT.
DR   PRINTS; PR01488; RTXTOXINA.
DR   PROSITE; PS00166; ENOYL_COA_HYDRATASE; 1.
SQ   SEQUENCE   953 AA;  102138 MW;  C41D9EBC1D799951 CRC64;

  Query Match             1.1%;  Score 10;  DB 2;  Length 953;
  Best Local Similarity   100.0%;  Pred. No. 1.7;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        265 LAQRVAAGLS 274
              ||||||||||
Db        283 LAQRVAAGLS 292

RESULT 26
Q9ETG5
ID   Q9ETG5      PRELIMINARY;      PRT;   953 AA.
AC   Q9ETG5;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Leukotoxin.
GN   LKTA.
OS   Pasteurella haemolytica.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC   Pasteurellaceae; Mannheimia.
OX   NCBI_TaxID=75985;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PH284, PH388, and PH8;
RA   Davies R.L., Whittam T.S., Selander R.K.;
RT   "Sequence diversity and molecular evolution of the leukotoxin (lktA)
RT   gene in bovine and ovine strains of Mannheimia (Pasteurella)
RT   haemolytica.";
RL   J. Bacteriol. 0:0-0(2001).
DR   EMBL; AF314507; AAG40291.1; -.
DR   EMBL; AF314504; AAG40288.1; -.
DR   EMBL; AF314506; AAG40290.1; -.
DR   InterPro; IPR001753; EnCoA_hydrtse.
DR   InterPro; IPR001343; Hemlysn_Ca_bind.
DR   InterPro; IPR003995; RtxA.
DR   Pfam; PF00353; hemolysinCabind; 5.
DR   Pfam; PF02382; RTX; 1.
DR   PRINTS; PR00313; CABNDNGRPT.
DR   PRINTS; PR01488; RTXTOXINA.
DR   PROSITE; PS00166; ENOYL_COA_HYDRATASE; 1.
SQ   SEQUENCE   953 AA;  101997 MW;  D593D6A577C3ADE9 CRC64;

  Query Match             1.1%;  Score 10;  DB 2;  Length 953;
  Best Local Similarity   100.0%;  Pred. No. 1.7;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        265 LAQRVAAGLS 274
              ||||||||||
Db        283 LAQRVAAGLS 292

RESULT 27
Q9EV23
ID   Q9EV23      PRELIMINARY;      PRT;   953 AA.
AC   Q9EV23;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Leukotoxin.
GN   LKTA.
OS   Mannheimia glucosida.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC   Pasteurellaceae; Mannheimia.
OX   NCBI_TaxID=85401;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PH290;
RX   MEDLINE=21101823; PubMed=11157953;
RA   Davies R.L., Whittam T.S., Selander R.K.;
RT   "Sequence Diversity and Molecular Evolution of the Leukotoxin (lktA)
RT   Gene in Bovine and Ovine Strains of Mannheimia (Pasteurella)
RT   haemolytica.";
RL   J. Bacteriol. 183:1394-1404(2001).
```

```
DR   EMBL; AF314522; AAG40306.1; -.
DR   InterPro; IPR001753; EnCoA_hydrtse.
DR   InterPro; IPR001343; Hemlysn_Ca_bind.
DR   InterPro; IPR003995; RtxA.
DR   Pfam; PF00353; hemolysinCabind; 5.
DR   Pfam; PF02382; RTX; 1.
DR   PRINTS; PR00313; CABNDNGRPT.
DR   PRINTS; PR01488; RTXTOXINA.
DR   PROSITE; PS00166; ENOYL_COA_HYDRATASE; 1.
SQ   SEQUENCE   953 AA;  102150 MW;  D99C36DA595B1624 CRC64;

  Query Match             1.1%;  Score 10;  DB 2;  Length 953;
  Best Local Similarity   100.0%;  Pred. No. 1.7;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        265 LAQRVAAGLS 274
              ||||||||||
Db        283 LAQRVAAGLS 292


RESULT 28
Q9EV33
ID   Q9EV33      PRELIMINARY;      PRT;   953 AA.
AC   Q9EV33;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Leukotoxin.
GN   LKTA.
OS   Pasteurella haemolytica.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC   Pasteurellaceae; Mannheimia.
OX   NCBI_TaxID=75985;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PH66;
RX   MEDLINE=21101823; PubMed=11157953;
RA   Davies R.L., Whittam T.S., Selander R.K.;
RT   "Sequence Diversity and Molecular Evolution of the Leukotoxin (lktA)
RT   Gene in Bovine and Ovine Strains of Mannheimia (Pasteurella)
RT   haemolytica.";
RL   J. Bacteriol. 183:1394-1404(2001).
DR   EMBL; AF314508; AAG40292.1; -.
DR   InterPro; IPR001753; EnCoA_hydrtse.
DR   InterPro; IPR001343; Hemlysn_Ca_bind.
DR   InterPro; IPR003995; RtxA.
DR   Pfam; PF00353; hemolysinCabind; 5.
DR   Pfam; PF02382; RTX; 1.
DR   PRINTS; PR00313; CABNDNGRPT.
DR   PRINTS; PR01488; RTXTOXINA.
DR   PROSITE; PS00166; ENOYL_COA_HYDRATASE; 1.
SQ   SEQUENCE   953 AA;  102132 MW;  4138AB5FAE2843B3 CRC64;

  Query Match             1.1%;  Score 10;  DB 2;  Length 953;
  Best Local Similarity   100.0%;  Pred. No. 1.7;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        265 LAQRVAAGLS 274
              ||||||||||
Db        283 LAQRVAAGLS 292


RESULT 29
Q9EV34
ID   Q9EV34      PRELIMINARY;      PRT;   953 AA.
AC   Q9EV34;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Leukotoxin.
GN   LKTA.
OS   Pasteurella haemolytica.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC   Pasteurellaceae; Mannheimia.
OX   NCBI_TaxID=75985;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PH56;
RX   MEDLINE=21101823; PubMed=11157953;
RA   Davies R.L., Whittam T.S., Selander R.K.;
RT   "Sequence Diversity and Molecular Evolution of the Leukotoxin (lktA)
RT   Gene in Bovine and Ovine Strains of Mannheimia (Pasteurella)
RT   haemolytica.";
RL   J. Bacteriol. 183:1394-1404(2001).
DR   EMBL; AF314505; AAG40289.1; -.
DR   InterPro; IPR001753; EnCoA_hydrtse.
DR   InterPro; IPR001343; Hemlysn_Ca_bind.
DR   InterPro; IPR003995; RtxA.
DR   Pfam; PF00353; hemolysinCabind; 5.
DR   Pfam; PF02382; RTX; 1.
DR   PRINTS; PR00313; CABNDNGRPT.
DR   PRINTS; PR01488; RTXTOXINA.
DR   PROSITE; PS00166; ENOYL_COA_HYDRATASE; 1.
SQ   SEQUENCE   953 AA;  102010 MW;  DE48B28EE0EB09FB CRC64;

  Query Match             1.1%;  Score 10;  DB 2;  Length 953;
  Best Local Similarity   100.0%;  Pred. No. 1.7;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        265 LAQRVAAGLS 274
              ||||||||||
Db        283 LAQRVAAGLS 292


RESULT 30
Q9EV32
ID   Q9EV32      PRELIMINARY;      PRT;   953 AA.
AC   Q9EV32;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Leukotoxin.
GN   LKTA.
OS   Pasteurella haemolytica.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC   Pasteurellaceae; Mannheimia.
OX   NCBI_TaxID=75985;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PH706;
RX   MEDLINE=21101823; PubMed=11157953;
RA   Davies R.L., Whittam T.S., Selander R.K.;
RT   "Sequence Diversity and Molecular Evolution of the Leukotoxin (lktA)
RT   Gene in Bovine and Ovine Strains of Mannheimia (Pasteurella)
RT   haemolytica.";
RL   J. Bacteriol. 183:1394-1404(2001).
DR   EMBL; AF314509; AAG40293.1; -.
DR   InterPro; IPR001753; EnCoA_hydrtse.
DR   InterPro; IPR001343; Hemlysn_Ca_bind.
DR   InterPro; IPR003995; RtxA.
DR   Pfam; PF00353; hemolysinCabind; 5.
DR   Pfam; PF02382; RTX; 1.
DR   PRINTS; PR00313; CABNDNGRPT.
DR   PRINTS; PR01488; RTXTOXINA.
DR   PROSITE; PS00166; ENOYL_COA_HYDRATASE; 1.
SQ   SEQUENCE   953 AA;  102078 MW;  EF425243C8741EE4 CRC64;

  Query Match             1.1%;  Score 10;  DB 2;  Length 953;
  Best Local Similarity   100.0%;  Pred. No. 1.7;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        265 LAQRVAAGLS 274
              ||||||||||
Db        283 LAQRVAAGLS 292
```

```
RESULT 31
Q9EV30
ID   Q9EV30      PRELIMINARY;      PRT;   953 AA.
AC   Q9EV30;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Leukotoxin.
GN   LKTA.
OS   Pasteurella haemolytica.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC   Pasteurellaceae; Mannheimia.
OX   NCBI_TaxID=75985;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PH196;
RX   MEDLINE=21101823; PubMed=11157953;
RA   Davies R.L., Whittam T.S., Selander R.K.;
RT   "Sequence Diversity and Molecular Evolution of the Leukotoxin (lktA)
RT   Gene in Bovine and Ovine Strains of Mannheimia (Pasteurella)
RT   haemolytica.";
RL   J. Bacteriol. 183:1394-1404(2001).
DR   EMBL; AF314512; AAG40296.1; -.
DR   InterPro; IPR001753; EnCoA_hydrtse.
DR   InterPro; IPR001343; Hemlysn_Ca_bind.
DR   InterPro; IPR003995; RtxA.
DR   Pfam; PF00353; hemolysinCabind; 5.
DR   Pfam; PF02382; RTX; 1.
DR   PRINTS; PR00313; CABNDNGRPT.
DR   PRINTS; PR01488; RTXTOXINA.
DR   PROSITE; PS00166; ENOYL_COA_HYDRATASE; 1.
SQ   SEQUENCE   953 AA;  102043 MW;  4E8F11490479A69A CRC64;

  Query Match             1.1%;  Score 10;  DB 2;  Length 953;
  Best Local Similarity   100.0%;  Pred. No. 1.7;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        265 LAQRVAAGLS 274
              ||||||||||
Db        283 LAQRVAAGLS 292

RESULT 32
Q9EV28
ID   Q9EV28      PRELIMINARY;      PRT;   953 AA.
AC   Q9EV28;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Leukotoxin (LktA).
GN   LKTA.
OS   Pasteurella haemolytica.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC   Pasteurellaceae; Mannheimia.
OX   NCBI_TaxID=75985;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PH292, and PH296;
RX   MEDLINE=21101823; PubMed=11157953;
RA   Davies R.L., Whittam T.S., Selander R.K.;
RT   "Sequence Diversity and Molecular Evolution of the Leukotoxin (lktA)
RT   Gene in Bovine and Ovine Strains of Mannheimia (Pasteurella)
RT   haemolytica.";
RL   J. Bacteriol. 183:1394-1404(2001).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PH296;
RA   Davies R.L., Campbell S., Whittam T.S.;
RT   "Mosaic structure and molecular evolution of the leukotoxin operon
RT   (lktCABD) of Mannheimia (pasteurella) haemolytica, Mannheimia
RT   glucosida and Pasteurella trehalosi.";
RL   J. Bacteriol. 0:0-0(2001).
DR   EMBL; AF314515; AAG40299.1; -.
DR   EMBL; AF414141; AAL13281.1; -.
DR   InterPro; IPR001753; EnCoA_hydrtse.
DR   InterPro; IPR001343; Hemlysn_Ca_bind.
DR   InterPro; IPR003995; RtxA.
DR   Pfam; PF00353; hemolysinCabind; 5.
DR   Pfam; PF02382; RTX; 1.
DR   PRINTS; PR00313; CABNDNGRPT.
DR   PRINTS; PR01488; RTXTOXINA.
DR   PROSITE; PS00166; ENOYL_COA_HYDRATASE; 1.
SQ   SEQUENCE   953 AA;  102218 MW;  04AB1715B819E571 CRC64;

  Query Match             1.1%;  Score 10;  DB 2;  Length 953;
  Best Local Similarity   100.0%;  Pred. No. 1.7;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        265 LAQRVAAGLS 274
              ||||||||||
Db        283 LAQRVAAGLS 292

RESULT 33
Q9EV26
ID   Q9EV26      PRELIMINARY;      PRT;   953 AA.
AC   Q9EV26;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Leukotoxin.
GN   LKTA.
OS   Mannheimia glucosida.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC   Pasteurellaceae; Mannheimia.
OX   NCBI_TaxID=85401;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PH240;
RX   MEDLINE=21101823; PubMed=11157953;
RA   Davies R.L., Whittam T.S., Selander R.K.;
RT   "Sequence Diversity and Molecular Evolution of the Leukotoxin (lktA)
RT   Gene in Bovine and Ovine Strains of Mannheimia (Pasteurella)
RT   haemolytica.";
RL   J. Bacteriol. 183:1394-1404(2001).
DR   EMBL; AF314519; AAG40303.1; -.
DR   InterPro; IPR001753; EnCoA_hydrtse.
DR   InterPro; IPR001343; Hemlysn_Ca_bind.
DR   InterPro; IPR003995; RtxA.
DR   Pfam; PF00353; hemolysinCabind; 5.
DR   Pfam; PF02382; RTX; 1.
DR   PRINTS; PR00313; CABNDNGRPT.
DR   PRINTS; PR01488; RTXTOXINA.
DR   PROSITE; PS00166; ENOYL_COA_HYDRATASE; 1.
SQ   SEQUENCE   953 AA;  102236 MW;  AC5874B3B75D1C98 CRC64;

  Query Match             1.1%;  Score 10;  DB 2;  Length 953;
  Best Local Similarity   100.0%;  Pred. No. 1.7;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        265 LAQRVAAGLS 274
              ||||||||||
Db        283 LAQRVAAGLS 292

RESULT 34
Q9EUE1
ID   Q9EUE1      PRELIMINARY;      PRT;   954 AA.
AC   Q9EUE1;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Leukotoxin.
GN   LKTA.
OS   Pasteurella haemolytica.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
```

```
OC   Pasteurellaceae; Mannheimia.
OX   NCBI_TaxID=75985;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PH202, and PH494;
RA   Davies R.L., Whittam T.S., Selander R.K.;
RT   "Sequence diversity and molecular evolution of the leukotoxin (lktA)
RT   gene in bovine and ovine strains of Mannheimia (Pasteurella)
RT   haemolytica.";
RL   J. Bacteriol. 0:0-0(2001).
DR   EMBL; AF314513; AAG40297.1; -.
DR   EMBL; AF314511; AAG40295.1; -.
DR   InterPro; IPR001753; EnCoA_hydrtse.
DR   InterPro; IPR001343; Hemlysn_Ca_bind.
DR   InterPro; IPR003995; RtxA.
DR   Pfam; PF00353; hemolysinCabind; 5.
DR   Pfam; PF02382; RTX; 1.
DR   PRINTS; PR00313; CABNDNGRPT.
DR   PRINTS; PR01488; RTXTOXINA.
DR   PROSITE; PS00166; ENOYL_COA_HYDRATASE; 1.
SQ   SEQUENCE   954 AA;  102410 MW;  02A60E52411711A2 CRC64;

  Query Match             1.1%;  Score 10;  DB 2;  Length 954;
  Best Local Similarity   100.0%;  Pred. No. 1.7;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        265 LAQRVAAGLS 274
              ||||||||||
Db        283 LAQRVAAGLS 292


RESULT 35
Q9EUD4
ID   Q9EUD4      PRELIMINARY;      PRT;   955 AA.
AC   Q9EUD4;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Leukotoxin.
GN   LKTA.
OS   Pasteurella trehalosi.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC   Pasteurellaceae; Pasteurella.
OX   NCBI_TaxID=47735;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PH254, PH246, and PH252;
RA   Davies R.L., Whittam T.S., Selander R.K.;
RT   "Sequence diversity and molecular evolution of the leukotoxin (lktA)
RT   gene in bovine and ovine strains of Mannheimia (Pasteurella)
RT   haemolytica.";
RL   J. Bacteriol. 0:0-0(2001).
DR   EMBL; AF314525; AAG40309.1; -.
DR   EMBL; AF314523; AAG40307.1; -.
DR   EMBL; AF314524; AAG40308.1; -.
DR   InterPro; IPR001753; EnCoA_hydrtse.
DR   InterPro; IPR001343; Hemlysn_Ca_bind.
DR   InterPro; IPR003995; RtxA.
DR   Pfam; PF00353; hemolysinCabind; 5.
DR   Pfam; PF02382; RTX; 1.
DR   PRINTS; PR00313; CABNDNGRPT.
DR   PRINTS; PR01488; RTXTOXINA.
DR   PROSITE; PS00166; ENOYL_COA_HYDRATASE; 1.
SQ   SEQUENCE   955 AA;  102127 MW;  074B2E4ADFFF57AA CRC64;

  Query Match             1.1%;  Score 10;  DB 2;  Length 955;
  Best Local Similarity   100.0%;  Pred. No. 1.7;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        265 LAQRVAAGLS 274
              ||||||||||
Db        285 LAQRVAAGLS 294


RESULT 36
Q9EV22
ID   Q9EV22      PRELIMINARY;      PRT;   955 AA.
AC   Q9EV22;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Leukotoxin.
GN   LKTA.
OS   Pasteurella trehalosi.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC   Pasteurellaceae; Pasteurella.
OX   NCBI_TaxID=47735;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PH68;
RX   MEDLINE=21101823; PubMed=11157953;
RA   Davies R.L., Whittam T.S., Selander R.K.;
RT   "Sequence Diversity and Molecular Evolution of the Leukotoxin (lktA)
RT   Gene in Bovine and Ovine Strains of Mannheimia (Pasteurella)
RT   haemolytica.";
RL   J. Bacteriol. 183:1394-1404(2001).
DR   EMBL; AF314526; AAG40310.1; -.
DR   InterPro; IPR001753; EnCoA_hydrtse.
DR   InterPro; IPR001343; Hemlysn_Ca_bind.
DR   InterPro; IPR003995; RtxA.
DR   Pfam; PF00353; hemolysinCabind; 5.
DR   Pfam; PF02382; RTX; 1.
DR   PRINTS; PR00313; CABNDNGRPT.
DR   PRINTS; PR01488; RTXTOXINA.
DR   PROSITE; PS00166; ENOYL_COA_HYDRATASE; 1.
SQ   SEQUENCE   955 AA;  102069 MW;  FDCF2E99B8DC394F CRC64;

  Query Match             1.1%;  Score 10;  DB 2;  Length 955;
  Best Local Similarity   100.0%;  Pred. No. 1.7;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        265 LAQRVAAGLS 274
              ||||||||||
Db        285 LAQRVAAGLS 294


RESULT 37
Q44494
ID   Q44494      PRELIMINARY;      PRT;  1403 AA.
AC   Q44494;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Mannuronan C-5-epimerase.
GN   ALGE1.
OS   Azotobacter vinelandii.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC   Pseudomonadaceae; Azotobacter.
OX   NCBI_TaxID=354;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=E;
RX   MEDLINE=94245609; PubMed=8188585;
RA   Ertesvag H., Doseth B., Larsen B., Skjak-Braek G., Valla S.;
RT   "Cloning and expression of an Azotobacter vinelandii mannuronan C-5-
RT   epimerase gene.";
RL   J. Bacteriol. 176:2846-2853(1994).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=E;
RX   MEDLINE=96065700; PubMed=7476166;
RA   Ertesvag H., Hoidal H.K., Hals I.K., Rian A., Doseth B., Valla S.;
RT   "A family of modular type mannuronan C-5-epimerase genes controls
RT   alginate structure in Azotobacter vinelandii.";
RL   Mol. Microbiol. 16:719-731(1995).
DR   EMBL; L39096; AAA87311.1; -.
```

```
DR   InterPro; IPR006633; CASH.
DR   InterPro; IPR001343; Hemlysn_Ca_bind.
DR   InterPro; IPR006626; PbH1.
DR   Pfam; PF00353; hemolysinCabind; 12.
DR   SMART; SM00722; CASH; 4.
DR   SMART; SM00710; PbH1; 15.
SQ   SEQUENCE   1403 AA;  147168 MW;  4E843AB0A366A95C CRC64;

  Query Match             1.1%;  Score 10;  DB 2;  Length 1403;
  Best Local Similarity   100.0%;  Pred. No. 2.5;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        714 GDDLLDGGAG 723
              ||||||||||
Db       1257 GDDLLDGGAG 1266

RESULT 38
Q9VL42
ID   Q9VL42       PRELIMINARY;      PRT;  1424 AA.
AC   Q9VL42;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   CG31714 protein.
GN   CG5833 OR CG13132.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Berkeley;
RX   MEDLINE=20196006; PubMed=10731132;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Celniker S.E., Adams M.D., Kronmiller B., Wan K.H., Holt R.A.,
RA   Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,
RA   Banzon J., An H., Baldwin D., Banzon J., Beeson K.Y., Busam D.A.,
RA   Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,
RA   Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,
RA   Ferriera S., Frise E., Galle R.F., Garg N.S., George R.A.,
RA   Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,
RA   Ibegwam C., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,
RA   McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,
RA   Pacleb J., Paragas V., Park S., Patel S., Pfeiffer B.,
RA   Phouanenavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,
RA   Stapleton M., Strong R., Svirskas R., Tector C., Tyler D.,
RA   Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;
RT   "Sequencing of Drosophila melanogaster genome.";
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Misra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,
RA   Hradecky P., Huang Y., Kaminker J.S., Prochnik S.E., Smith C.D.,
RA   Tupy J.L., Bergman C., Berman B., Carlson J.W., Celniker S.E.,
RA   Clamp M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,
RA   Kronmiller B., Marshall B., Millburn G., Richter J., Russo S.,
RA   Searle S.M.J., Smith E., Shu S., Smutniak F., Whitfield E.,
RA   Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.;
RT   "Annotation of Drosophila melanogaster genome.";
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE FROM N.A.
RA   Adams M.D., Celniker S.E., Gibbs R.A., Rubin G.M., Venter C.J.;
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [5]
RP   SEQUENCE FROM N.A.
RA   FlyBase;
RL   Submitted (SEP-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AE003627; AAF52856.2; -.
DR   FlyBase; FBgn0032180; CG5833.
DR   InterPro; IPR001879; hormn_receptor.
DR   InterPro; IPR003599; Ig.
DR   InterPro; IPR007110; Ig-like.
DR   Pfam; PF02793; HRM; 1.
DR   SMART; SM00409; IG; 1.
DR   PROSITE; PS50227; G_PROTEIN_RECEP_F2_3; 1.
DR   PROSITE; PS50835; IG_LIKE; 1.
SQ   SEQUENCE   1424 AA;  152768 MW;  E8E39E276FD7414E CRC64;

  Query Match             1.1%;  Score 10;  DB 5;  Length 1424;
  Best Local Similarity   100.0%;  Pred. No. 2.5;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        353 AGVSAAAVGS 362
              ||||||||||
Db       1106 AGVSAAAVGS 1115


RESULT 39
Q8YKJ3
ID   Q8YKJ3       PRELIMINARY;      PRT;  4936 AA.
AC   Q8YKJ3;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Hypothetical protein Alr7304.
GN   ALR7304.
OS   Anabaena sp. (strain PCC 7120).
OG   Plasmid pCC7120alpha.
OC   Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
OX   NCBI_TaxID=103690;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21595285; PubMed=11759840;
RA   Kaneko T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,
```

```
RA   Watanabe A., Iriguchi M., Ishikawa A., Kawashima K., Kimura T.,
RA   Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
RA   Nakazaki N., Shimpo S., Sugimoto M., Takazawa M., Yamada M.,
RA   Yasuda M., Tabata S.;
RT   "Complete genomic sequence of the filamentous nitrogen-fixing
RT   cyanobacterium Anabaena sp. strain PCC 7120.";
RL   DNA Res. 8:205-213(2001).
DR   EMBL; AP003600; BAB78388.1; -.
DR   InterPro; IPR006644; Cadg.
DR   InterPro; IPR003644; Calx_beta.
DR   InterPro; IPR002048; EF-hand.
DR   InterPro; IPR001343; Hemlysn_Ca_bind.
DR   InterPro; IPR000209; Peptidase_S8.
DR   Pfam; PF03160; Calx-beta; 5.
DR   Pfam; PF00353; hemolysinCabind; 25.
DR   Pfam; PF00082; Peptidase_S8; 1.
DR   PRINTS; PR00313; CABNDNGRPT.
DR   PRINTS; PR00723; SUBTILISIN.
DR   SMART; SM00736; CADG; 2.
DR   SMART; SM00237; Calx_beta; 6.
DR   PROSITE; PS00018; EF_HAND; 1.
DR   PROSITE; PS00330; HEMOLYSIN_CALCIUM; 2.
DR   PROSITE; PS00136; SUBTILASE_ASP; 1.
KW   Plasmid; Hypothetical protein; Complete proteome.
SQ   SEQUENCE   4936 AA;  519407 MW;  FB5A8323CB29C828 CRC64;

  Query Match             1.1%;  Score 10;  DB 16;  Length 4936;
  Best Local Similarity   100.0%;  Pred. No. 8;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        752 DVLNGGAGND 761
              ||||||||||
Db       4296 DVLNGGAGND 4305


RESULT 40
Q9RDA8
ID   Q9RDA8      PRELIMINARY;      PRT;   259 AA.
AC   Q9RDA8;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Hypothetical protein SCO2843.
GN   SCO2843 OR SCE20.17C.
OS   Streptomyces coelicolor.
OC   Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC   Streptomycineae; Streptomycetaceae; Streptomyces.
OX   NCBI_TaxID=1902;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A3(2);
RA   Seeger K.J., Harris D.;
RL   Submitted (JAN-2000) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A3(2);
RA   Thomson N.R., Parkhill J., Barrell B.G., Rajandream M.A.;
RL   Submitted (JAN-2000) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A3(2);
RX   MEDLINE=97000351; PubMed=8843436;
RA   Redenbach M., Kieser H.M., Denapaite D., Eichner A., Cullum J.,
RA   Kinashi H., Hopwood D.A.;
RT   "A set of ordered cosmids and a detailed genetic and physical map for
RT   the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
RL   Mol. Microbiol. 21:77-96(1996).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A3(2) / M145;
RX   MEDLINE=21996410; PubMed=12000953;
RA   Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
RA   Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA   Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA   Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA   Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neil S.,
RA   Rabbinowitsch E., Rajandream M.A., Rutherford K., Rutter S.,
RA   Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA   Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,
RA   Hopwood D.A.;
RT   "Complete genome sequence of the model actinomycete Streptomyces
RT   coelicolor A3(2).";
RL   Nature 417:141-147(2002).
DR   EMBL; AL939114; CAB65573.1; -.
DR   InterPro; IPR006357; HAD_SF_IIA.
DR   InterPro; IPR005834; Hydrolase.
DR   Pfam; PF00702; Hydrolase; 1.
DR   TIGRFAMs; TIGR01460; HAD-SF-IIA; 1.
KW   Hypothetical protein; Complete proteome.
SQ   SEQUENCE   259 AA;  28059 MW;  997D2583D4F7835A CRC64;

  Query Match             1.0%;  Score 9;  DB 16;  Length 259;
  Best Local Similarity   100.0%;  Pred. No. 5.5;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        276 TGAVAALIT 284
              |||||||||
Db        164 TGAVAALIT 172


Search completed: February 17, 2004, 10:21:24
Job time : 85 secs
```

```
OM protein - protein search, using sw model

Run on:         February 17, 2004, 10:09:49 ; Search time 21 Seconds
                                              (without alignments)
                                              1867.722 Million cell updates/sec

Title:          US-10-069-799-5
Perfect score:  4647
Sequence:       1 MSNINVIKSNIQAGLNSTKS..........SSNALQPITQPTQGILAPSV 927

Scoring table:  BLOSUM62
                Gapop 10.0 , Gapext 0.5

Searched:       328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters:      328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       Issued_Patents_AA:*
                1:  /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
                2:  /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
                3:  /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
                4:  /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
                5:  /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
                6:  /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

     Pred. No. is the number of results predicted by chance to have a
     score greater than or equal to the score of the result being printed,
     and is derived by analysis of the total score distribution.
```

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2334	50.2	1098	1	US-07-777-715-7	Sequence 7, Appli
2	2334	50.2	1098	1	US-08-170-126-2	Sequence 2, Appli
3	2334	50.2	1098	3	US-08-954-418-2	Sequence 2, Appli
4	2326.5	50.1	924	3	US-08-619-812-8	Sequence 8, Appli
5	2326.5	50.1	926	1	US-07-908-253-2	Sequence 2, Appli
6	2326.5	50.1	926	1	US-08-455-970A-2	Sequence 2, Appli
7	2326.5	50.1	926	1	US-08-387-156-6	Sequence 6, Appli
8	2326.5	50.1	926	2	US-08-694-865-6	Sequence 6, Appli
9	2326.5	50.1	926	2	US-08-878-748-6	Sequence 6, Appli
10	2326.5	50.1	926	2	US-08-535-837-2	Sequence 2, Appli
11	2326.5	50.1	926	3	US-09-124-491-6	Sequence 6, Appli
12	2326.5	50.1	926	4	US-09-383-912-6	Sequence 6, Appli
13	2326.5	50.1	926	6	5476657-3	Patent No. 5476657
14	2326.5	50.1	936	1	US-08-455-970A-12	Sequence 12, Appl
15	2326.5	50.1	943	1	US-08-455-970A-10	Sequence 10, Appl
16	2326.5	50.1	951	1	US-08-455-970A-14	Sequence 14, Appl
17	2326.5	50.1	977	1	US-08-387-156-8	Sequence 8, Appli
18	2326.5	50.1	977	2	US-08-694-865-8	Sequence 8, Appli
19	2326.5	50.1	977	2	US-08-878-748-8	Sequence 8, Appli
20	2326.5	50.1	977	3	US-09-124-491-8	Sequence 8, Appli
21	2326.5	50.1	977	4	US-09-383-912-8	Sequence 8, Appli
22	2326.5	50.1	1069	1	US-07-777-715-9	Sequence 9, Appli
23	2326.5	50.1	1069	1	US-08-170-126-4	Sequence 4, Appli
24	2326.5	50.1	1069	3	US-08-954-418-4	Sequence 4, Appli
25	2297.5	49.4	934	1	US-08-215-805A-80	Sequence 80, Appl
26	2236.5	48.1	956	3	US-08-772-270A-8	Sequence 8, Appli
27	2236.5	48.1	956	4	US-09-062-126-8	Sequence 8, Appli
28	1926	41.4	1053	4	US-09-062-126-10	Sequence 10, Appl
29	1924	41.4	1049	3	US-08-772-270A-11	Sequence 11, Appl
30	1924	41.4	1244	5	PCT-US93-10500-2	Sequence 2, Appli
31	1763.5	37.9	1022	3	US-08-772-270A-2	Sequence 2, Appli
32	1751.5	37.7	1013	4	US-09-062-126-3	Sequence 3, Appli
33	1376.5	29.6	758	1	US-08-258-188-2	Sequence 2, Appli
34	1376.5	29.6	758	1	US-08-526-813-2	Sequence 2, Appli
35	1376.5	29.6	758	5	PCT-US95-08554-2	Sequence 2, Appli
36	904.5	19.5	544	1	US-08-387-156-10	Sequence 10, Appl
37	904.5	19.5	544	2	US-08-694-865-10	Sequence 10, Appl
38	904.5	19.5	544	2	US-08-878-748-10	Sequence 10, Appl
39	904.5	19.5	544	3	US-09-124-491-10	Sequence 10, Appl
40	904.5	19.5	544	4	US-09-383-912-10	Sequence 10, Appl
41	901	19.4	699	2	US-08-694-865-16	Sequence 16, Appl
42	901	19.4	699	3	US-09-124-491-16	Sequence 16, Appl
43	901	19.4	699	4	US-09-383-912-16	Sequence 16, Appl
44	897.5	19.3	490	4	US-09-252-149B-26	Sequence 26, Appl
45	896.5	19.3	1403	1	US-07-908-253-3	Sequence 3, Appli

ALIGNMENTS

```
RESULT 1
US-07-777-715-7
; Sequence 7, Application US/07777715
; Patent No. 5273889
;  GENERAL INFORMATION:
;    APPLICANT:  Potter, Andrew
;    APPLICANT:  Campos, Manuel
;    APPLICANT:  Hughes, Huw P.A.
;    TITLE OF INVENTION:  CYTOKINE-LEUKOTOXIN GENE FUSIONS AND
;    TITLE OF INVENTION:  USES THEREOF
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Morrison & Foerster
;      STREET:  545 Middlefield Road, Suite 200
;      CITY:  Menlo Park
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94025
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/777,715
;      FILING DATE:  19911016
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Robins, Roberta L.
;      REGISTRATION NUMBER:  33,208
;      REFERENCE/DOCKET NUMBER:  29310-2001320
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-327-7250
;      TELEFAX:  415-327-2951
;      TELEX:  706141
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1098 amino acids
;      TYPE:  AMINO ACID
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-07-777-715-7

  Query Match             50.2%;  Score 2334;  DB 1;  Length 1098;
  Best Local Similarity   50.1%;  Pred. No. 5.2e-174;
  Matches  465;  Conservative  175;  Mismatches  263;  Indels   26;  Gaps   13;

Qy          8 KSNIQAGLNSTKSGLKNLYLAIPKD--YDPQKGGTLNDFIKAADELGIARLAEEPNHTET 65
              :|  ||| :| |:| | : | ||::  || ::|  | | :|||:||||    || |:  |
Db        173 QSLTQAG-SSLKTGAKKIILYIPQNYQYDTEQGNGLQDLVKAAEELGIEVQREERNNIAT 231
```

```
Qy         66 AKKSVDTVNQFLSLTQTGIAISATKLEKFLQKHSTNKLAKGLDSVENIDRKLGKASNVLS 125
Db        232 AQTSLGTIQTAIGLTERGIVLSAPQIDKLLQK---TKAGQALGSAESIVQNANKAKTVLS 288

Qy        126 TLSSFLGTALAGIELDSLIKKGDAAPDALAKASIDLINEIIGNLSQSTQTIEAFSSQLAK 185
Db        289 GIQSILGSVLAGMDLDEAL-QNNSNQHALAKAGLELTNSLIENIANSVKTLDEFGEQISQ 347

Qy        186 LGSTISQAKGFSNIGNKLQNL-NFSKTNLGLEIITGLLSGISAGFALADKNASTGKKVAA 244
Db        348 FGSKLQNIKGLGTLGDKLKNIGGLDKAGLGLDVISGLLSGATAALVLADKNASTAKKVGA 407

Qy        245 GFELSNQVIGNVTKAISSYVLAQRVAAGLSTTGAVAALITSSIMLAISPLAFMNAADKFN 304
Db        408 GFELANQVVGNITKAVSSYILAQRVAAGLSSTGPVAALIASTVSLAISPLAFAGIADKFN 467

Qy        305 HANALDEFAKQFRKFGYDGDHLLAEYQRGVGTIEASLTTISTALGAVSAGVSAAAVGSAV 364
Db        468 HAKSLESYAERFKKLGYDGDNLLAEYQRGTGTIDASVTAINTALAAIAGGVSAAAAGSVI 527

Qy        365 GTPIALLVAGVTGLISGILEASKQAMFESVANRLQGKILEWEKQNGGQNYFDKGYDSRYA 424
Db        528 ASPIALLVSGITGVISTILQYSKQAMFEHVANKIHNKIVEWEKNNHGKNYFENGYDARYL 587

Qy        425 AYLANNLKFLSELNKELEAERVIAITQQRWDNNIGELAGITKLGERIKSGKAYADAFEDG 484
Db        588 ANLQDNMKFLLNLNKELQAERVIAITQQQWDNNIGDLAGISRLGEKVLSGKAYVDAFEEG 647

Qy        485 KKVEAGSNITLDAKTGIIDISNSNGKKTQALHFTSPLLTAGTESRERLTNGKYSYINKLK 544
Db        648 KHIKADKLVQLDSANGIIDVSNSGKAKTQHILFRTPLLTPGTEHRERVQTGKYEYITKLN 707

Qy        545 FGRVKNWQVTDGEASSKLDFSKVIQRV-------AETEGTDEIGLIVNAKAGNDDIFVGQ 597
Db        708 INRVDSWKITDGAASSTFDLTNVVQRIGIELDNAGNVTKTKETKIIAKLGEGDDNVFVGS 767

Qy        598 GKMNIDGGDGHDRVFYSKDGGFGNITVDGTSATEAGSYTVNRKVARGDIYHEVVKRQETK 657
Db        768 GTTEIDGGEGYDRVHYSR-GNYGALTIDATKETEQGSYTVNRFVETGKALHEVTSTHTAL 826

Qy        658 VGKRTETIQYRDYELRKVGYGYQSTDNLKSVEEVIGSQFNDVFKGSKFNDIFHSGEGDDL 717
Db        827 VGNREEKIEYR-HSNNQHHAGYYTKDTLKAVEEIIGTSHNDIFKGSKFNDAFNGGDGVDT 885

Qy        718 LDGGAGDDRLFGGKGNDRLSGDEGDDLLDGGSGDDVLNGGAGNDVYIFRKGDGNDTLYDG 777
Db        886 IDGNDGNDRLFGGKGDDILDGGNGDDFIDGGKGNDLLHGGKGDDIFVHRKGDGNDIITDS 945

Qy        778 TGNDKLAFADANISDIMIERTKEGIIVKRNDHSGSINIPRWY----ITSNLQNYQSNKTD 833
Db        946 DGNDKLSFSDSNLKDLTFEKVKHNLVI-TNSKKEKVTIQNWFREADFAKEVPNYKATK-D 1003

Qy        834 HKIEQLIGKDGSYITSDQIDKILQDKKDGTVITSQELKKLADENKSQKLSASDIASSLNK 893
Db       1004 EKIEEIIGQNGERITSKQVDDLI--AKGNGKITQDELSKVVDNYELLKHS-KNVTNSLDK 1060

Qy        894 LVGSMALFGTANSVSSNALQPITQPTQGI 922
Db       1061 LISSVSAFTSSNDSRNVLVAPTSMLDQSL 1089
```

RESULT 2
US-08-170-126-2

```
; Sequence 2, Application US/08170126
; Patent No. 5594107
;  GENERAL INFORMATION:
;    APPLICANT:  POTTER, ANDREW
;    APPLICANT:  CAMPOS, MANUEL
;    APPLICANT:  HUGHES, HUW P.A.
;    TITLE OF INVENTION:  CYTOKINE-CYTOTOXIN GENE FUSIONS AND USES
;    TITLE OF INVENTION:  THEREOF
;    NUMBER OF SEQUENCES:  6
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  REED & ROBINS
;      STREET:  635 BRYANT STREET
;      CITY:  PALO ALTO
;      STATE:  CALIFORNIA
;      COUNTRY:  UNITED STATES OF AMERICA
;      ZIP:  94301
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/170,126
;      FILING DATE:  20-DEC-1993
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/777,715
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/571,301
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  ROBINS, ROBERTA L.
;      REGISTRATION NUMBER:  33,208
;      REFERENCE/DOCKET NUMBER:  9000-0013.21
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 617-8999
;      TELEFAX:  (415) 327-3231
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1098 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-170-126-2

  Query Match             50.2%;  Score 2334;  DB 1;  Length 1098;
  Best Local Similarity   50.1%;  Pred. No. 5.2e-174;
  Matches  465;  Conservative  175;  Mismatches  263;  Indels   26;  Gaps   13;

Qy          8 KSNIQAGLNSTKSGLKNLYLAIPKD--YDPQKGGTLNDFIKAADELGIARLAEEPNHTET 65
Db        173 QSLTQAG-SSLKTGAKKIILYIPQNYQYDTEQGNGLQDLVKAAEELGIEVQREERNNIAT 231

Qy         66 AKKSVDTVNQFLSLTQTGIAISATKLEKFLQKHSTNKLAKGLDSVENIDRKLGKASNVLS 125
Db        232 AQTSLGTIQTAIGLTERGIVLSAPQIDKLLQK---TKAGQALGSAESIVQNANKAKTVLS 288

Qy        126 TLSSFLGTALAGIELDSLIKKGDAAPDALAKASIDLINEIIGNLSQSTQTIEAFSSQLAK 185
Db        289 GIQSILGSVLAGMDLDEAL-QNNSNQHALAKAGLELTNSLIENIANSVKTLDEFGEQISQ 347

Qy        186 LGSTISQAKGFSNIGNKLQNL-NFSKTNLGLEIITGLLSGISAGFALADKNASTGKKVAA 244
Db        348 FGSKLQNIKGLGTLGDKLKNIGGLDKAGLGLDVISGLLSGATAALVLADKNASTAKKVGA 407

Qy        245 GFELSNQVIGNVTKAISSYVLAQRVAAGLSTTGAVAALITSSIMLAISPLAFMNAADKFN 304
Db        408 GFELANQVVGNITKAVSSYILAQRVAAGLSSTGPVAALIASTVSLAISPLAFAGIADKFN 467

Qy        305 HANALDEFAKQFRKFGYDGDHLLAEYQRGVGTIEASLTTISTALGAVSAGVSAAAVGSAV 364
Db        468 HAKSLESYAERFKKLGYDGDNLLAEYQRGTGTIDASVTAINTALAAIAGGVSAAAAGSVI 527

Qy        365 GTPIALLVAGVTGLISGILEASKQAMFESVANRLQGKILEWEKQNGGQNYFDKGYDSRYA 424
Db        528 ASPIALLVSGITGVISTILQYSKQAMFEHVANKIHNKIVEWEKNNHGKNYFENGYDARYL 587

Qy        425 AYLANNLKFLSELNKELEAERVIAITQQRWDNNIGELAGITKLGERIKSGKAYADAFEDG 484
Db        588 ANLQDNMKFLLNLNKELQAERVIAITQQQWDNNIGDLAGISRLGEKVLSGKAYVDAFEEG 647

Qy        485 KKVEAGSNITLDAKTGIIDISNSNGKKTQALHFTSPLLTAGTESRERLTNGKYSYINKLK 544
```

```
Db        648 KHIKADKLVQLDSANGIIDVSNSGKAKTQHILFRTPLLTPGTEHRERVQTGKYEYITKLN 707

Qy        545 FGRVKNWQVTDGEASSKLDFSKVIQRV-------AETEGTDEIGLIVNAKAGNDDIFVGQ 597
                 || :|::||| |||  | : |:||:               | |  :|     |:|::|||
Db        708 INRVDSWKITDGAASSTFDLTNVVQRIGIELDNAGNVTKTKETKIIAKLGEGDDNVFVGS 767

Qy        598 GKMNIDGGDGHDRVFYSKDGGFGNITVDGTSATEAGSYTVNRKVARGDIYHEVVKRQETK 657
              |     ||||:|:||| ||: | :| :|:| |  || ||||||| |   |    |||
Db        768 GTTEIDGGEGYDRVHYSR-GNYGALTIDATKETEQGSYTVNRFVETGKALHEVTSTHTAL 826

Qy        658 VGKRTETIQYRDYELRKVGYGYQSTDNLKSVEEVIGSQFNDVFKGSKFNDIFHSGEGDDL 717
              || | | |:|| :   :      || : | ||:|||:||:  ||:||||||| |: |:| |
Db        827 VGNREEKIEYR-HSNNQHHAGYYTKDTLKAVEEIIGTSHNDIFKGSKFNDAFNGGDGVDT 885

Qy        718 LDGGAGDDRLFGGKGNDRLSGDEGDDLLDGGSGDDVLNGGAGNDVYIFRKGDGNDTLYDG 777
               :||  |:|||||||||:| |  |  |||:|||  |:|:|:|| |:|::: |||||| : |
Db        886 IDGNDGNDRLFGGKGDDILDGGNGDDFIDGGKGNDLLHGGKGDDIFVHRKGDGNDIITDS 945

Qy        778 TGNDKLAFADANISDIMIERTKEGIIVKRNDHSGSINIPRWY----ITSNLQNYQSNKTD 833
                |||||:|:|:|: |:  |:  |  :::  |       :  |  |:        : ||:: | |
Db        946 DGNDKLSFSDSNLKDLTFEKVKHNLVI-TNSKKEKVTIQNWFREADFAKEVPNYKATK-D 1003

Qy        834 HKIEQLIGKDGSYITSDQIDKILQDKKDGTVITSQELKKLADENKSQKLSASDIASSLNK 893
                |||::||::|  ||| |:| ::   |      ||  || |: |   :  | |   :: :||:|
Db       1004 EKIEEIIGQNGERITSKQVDDLI--AKGNGKITQDELSKVVDNYELLKHS-KNVTNSLDK 1060

Qy        894 LVGSMALFGTANSVSSNALQPITQPTQGI 922
              |: |:: | ::|   :  : | :    | :
Db       1061 LISSVSAFTSSNDSRNVLVAPTSMLDQSL 1089
```

RESULT 3
US-08-954-418-2
; Sequence 2, Application US/08954418
; Patent No. 6096320
; GENERAL INFORMATION:
;   APPLICANT: POTTER, ANDREW
;   APPLICANT: CAMPOS, MANUEL
;   APPLICANT: HUGHES, HUW P.A.
;   TITLE OF INVENTION: CYTOKINE-CYTOTOXIN GENE FUSIONS AND USES
;   TITLE OF INVENTION: THEREOF
;   NUMBER OF SEQUENCES: 6
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: REED & ROBINS
;   STREET: 635 BRYANT STREET
;   CITY: PALO ALTO
;   STATE: CALIFORNIA
;   COUNTRY: UNITED STATES OF AMERICA
;   ZIP: 94301
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.25
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/954,418
;   FILING DATE:
;   CLASSIFICATION:
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: 08/170,126
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US 07/571,301
;   ATTORNEY/AGENT INFORMATION:
;   NAME: ROBINS, ROBERTA L.
;   REGISTRATION NUMBER: 33,208
;   REFERENCE/DOCKET NUMBER: 9000-0013.21
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (415) 617-8999
;   TELEFAX: (415) 327-3231
; INFORMATION FOR SEQ ID NO: 2:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 1098 amino acids
;   TYPE: amino acid
;   TOPOLOGY: linear
;   MOLECULE TYPE: protein
US-08-954-418-2

```
  Query Match             50.2%;  Score 2334;  DB 3;  Length 1098;
  Best Local Similarity   50.1%;  Pred. No. 5.2e-174;
  Matches  465;  Conservative  175;  Mismatches  263;  Indels   26;  Gaps   13;

Qy          8 KSNIQAGLNSTKSGLKNLYLAIPKD--YDPQKGGTLNDFIKAADELGIARLAEEPNHTET 65
              :|  ||| :| |:| | : | ||::  || ::|  | | :|||:||||    || |:  |
Db        173 QSLTQAG-SSLKTGAKKIILYIPQNYQYDTEQGNGLQDLVKAAEELGIEVQREERNNIAT 231

Qy         66 AKKSVDTVNQFLSLTQTGIAISATKLEKFLQKHSTNKLAKGLDSVENIDRKLGKASNVLS 125
              |: |: |:    : ||: || :|| :::| |||      |   : | | |:| :   || |||
Db        232 AQTSLGTIQTAIGLTERGIVLSAPQIDKLLQK---TKAGQALGSAESIVQNANKAKTVLS 288

Qy        126 TLSSFLGTALAGIELDSLIKKGDAAPDALAKASIDLINEIIGNLSQSTQTIEAFSSQLAK 185
               : | ||: |||::||  : : ::   |||| ::| | :| |:: | :|:: |  |:::
Db        289 GIQSILGSVLAGMDLDEAL-QNNSNQHALAKAGLELTNSLIENIANSVKTLDEFGEQISQ 347

Qy        186 LGSTISQAKGFSNIGNKLQNL-NFSKTNLGLEIITGLLSGISAGFALADKNASTGKKVAA 244
               ||  :    ||      :|:||:|:     |    |||:|:|:||||  :|   ||||||| ||| |
Db        348 FGSKLQNIKGLGTLGDKLKNIGGLDKAGLGLDVISGLLSGATAALVLADKNASTAKKVGA 407

Qy        245 GFELSNQVIGNVTKAISSYVLAQRVAAGLSTTGAVAALITSSIMLAISPLAFMNAADKFN 304
              ||||:|||:||:|||:|||:||||||||:|| ||||| |::  ||||||     |||||
Db        408 GFELANQVVGNITKAVSSYILAQRVAAGLSSTGPVAALIASTVSLAISPLAFAGIADKFN 467

Qy        305 HANALDEFAKQFRKFGYDGDHLLAEYQRGVGTIEASLTTISTALGAVSAGVSAAAVGSAV 364
              || :|: :|::|:| |||||:||||||||  |||:|:| |:||| |:: ||||| || :
Db        468 HAKSLESYAERFKKLGYDGDNLLAEYQRGTGTIDASVTAINTALAAIAGGVSAAAAGSVI 527

Qy        365 GTPIALLVAGVTGLISGILEASKQAMFESVANRLQGKILEWEKQNGGQNYFDKGYDSRYA 424
               :|||||:|:||:|| ||: |||||| |||::  ||:|||| | |:|||: |||:|:|
Db        528 ASPIALLVSGITGVISTILQYSKQAMFEHVANKIHNKIVEWEKNNHGKNYFENGYDARYL 587

Qy        425 AYLANNLKFLSELNKELEAERVIAITQQRWDNNIGELAGITKLGERIKSGKAYADAFEDG 484
              | | :|:|||  |||||:|||||||||||:|||||||||::|||||:: |||||  ||||:|
Db        588 ANLQDNMKFLLNLNKELQAERVIAITQQQWDNNIGDLAGISRLGEKVLSGKAYVDAFEEG 647

Qy        485 KKVEAGSNITLDAKTGIIDISNSNGKKTQALHFTSPLLTAGTESRERLTNGKYSYINKLK 544
              | ::|     : ||:  |||:|||   ||| | :|||| ||| |||:  ||| || ||
Db        648 KHIKADKLVQLDSANGIIDVSNSGKAKTQHILFRTPLLTPGTEHRERVQTGKYEYITKLN 707

Qy        545 FGRVKNWQVTDGEASSKLDFSKVIQRV-------AETEGTDEIGLIVNAKAGNDDIFVGQ 597
                 || :|::||| |||  | : |:||:               | |  :|     |:|::|||
Db        708 INRVDSWKITDGAASSTFDLTNVVQRIGIELDNAGNVTKTKETKIIAKLGEGDDNVFVGS 767

Qy        598 GKMNIDGGDGHDRVFYSKDGGFGNITVDGTSATEAGSYTVNRKVARGDIYHEVVKRQETK 657
              |     ||||:|:||| ||: | :| :|:| |  || ||||||| |   |    |||
Db        768 GTTEIDGGEGYDRVHYSR-GNYGALTIDATKETEQGSYTVNRFVETGKALHEVTSTHTAL 826

Qy        658 VGKRTETIQYRDYELRKVGYGYQSTDNLKSVEEVIGSQFNDVFKGSKFNDIFHSGEGDDL 717
              || | | |:|| :   :      || : | ||:|||:||:  ||:||||||| |: |:| |
Db        827 VGNREEKIEYR-HSNNQHHAGYYTKDTLKAVEEIIGTSHNDIFKGSKFNDAFNGGDGVDT 885

Qy        718 LDGGAGDDRLFGGKGNDRLSGDEGDDLLDGGSGDDVLNGGAGNDVYIFRKGDGNDTLYDG 777
               :||  |:|||||||||:| |  |  |||:|||  |:|:|:|| |:|::: |||||| : |
Db        886 IDGNDGNDRLFGGKGDDILDGGNGDDFIDGGKGNDLLHGGKGDDIFVHRKGDGNDIITDS 945

Qy        778 TGNDKLAFADANISDIMIERTKEGIIVKRNDHSGSINIPRWY----ITSNLQNYQSNKTD 833
                |||||:|:|:|: |:  |:  |  :::  |       :  |  |:        : ||:: | |
Db        946 DGNDKLSFSDSNLKDLTFEKVKHNLVI-TNSKKEKVTIQNWFREADFAKEVPNYKATK-D 1003

Qy        834 HKIEQLIGKDGSYITSDQIDKILQDKKDGTVITSQELKKLADENKSQKLSASDIASSLNK 893
                |||::||::|  ||| |:| ::   |      ||  || |: |   :  | |   :: :||:|
Db       1004 EKIEEIIGQNGERITSKQVDDLI--AKGNGKITQDELSKVVDNYELLKHS-KNVTNSLDK 1060

Qy        894 LVGSMALFGTANSVSSNALQPITQPTQGI 922
              |: |:: | ::|   :  : | :    | :
Db       1061 LISSVSAFTSSNDSRNVLVAPTSMLDQSL 1089
```

```
RESULT 4
US-08-619-812-8
; Sequence 8, Application US/08619812
; Patent No. 6100066
;  GENERAL INFORMATION:
;    APPLICANT:  POTTER, ANDREW A.
;    APPLICANT:  THEISEN, MICHAEL
;    APPLICANT:  HARLAND, RICHARD J.
;    APPLICANT:  RIOUX, CLEMENT R.
;    TITLE OF INVENTION:  VACCINES FOR HAEMOPHILUS SOMNUS
;    NUMBER OF SEQUENCES:  10
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  REED & ROBINS
;      STREET:  635 BRYANT STREET
;      CITY:  PALO ALTO
;      STATE:  CALIFORNIA
;      COUNTRY:  UNITED STATES OF AMERICA
;      ZIP:  94301
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/619,812
;      FILING DATE:  15-MAR-1996
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/038,719
;      FILING DATE:  29-MAR-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  ROBINS, ROBERTA L.
;      REGISTRATION NUMBER:  33,208
;      REFERENCE/DOCKET NUMBER:  9000-0019.20
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 617-8999
;      TELEFAX:  (415) 327-3231
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  924 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-619-812-8

  Query Match             50.1%;  Score 2326.5;  DB 3;  Length 924;
  Best Local Similarity   50.0%;  Pred. No. 1.5e-173;
  Matches  461;  Conservative  174;  Mismatches  262;  Indels   25;  Gaps   12;

Qy         15 LNSTKSGLKNLYLAIPKD--YDPQKGGTLNDFIKAADELGIARLAEEPNHTETAKKSVDT 72
              |:  |:| | :   ||::  || ::|  | | :|||:||||    || |:  ||: |: |
Db          7 LSFPKTGAKKIILYIPQNYQYDTEQGNGLQDLVKAAEELGIEVQREERNNIATAQTSLGT 66

Qy         73 VNQFLSLTQTGIAISATKLEKFLQKHSTNKLAKGLDSVENIDRKLGKASNVLSTLSSFLG 132
              :   : ||: || :|| :::| |||    |  : | | |:| :   ||  ||| : | ||
Db         67 IQTAIGLTERGIVLSAPQIDKLLQK---TKAGQALGSAESIVQNANKAKTVLSGIQSILG 123

Qy        133 TALAGIELDSLIKKGDAAPDALAKASIDLINEIIGNLSQSTQTIEAFSSQLAKLGSTISQ 192
              : |||::||  : : ::   ||||| ::| | :| |:: | :|:: |  |::: || :  
Db        124 SVLAGMDLDEAL-QNNSNQHALAKAGLELTNSLIENIANSVKTLDEFGEQISQFGSKLQN 182

Qy        193 AKGFSNIGNKLQNL-NFSKTNLGLEIITGLLSGISAGFALADKNASTGKKVAAGFELSNQ 251
               ||    :|:|||:|:    |   |||::|:||||| :|   ||||||| ||| |||||:||
Db        183 IKGLGTLGDKLKNIGGLDKAGLGLDVISGLLSGATAALVLADKNASTAKKVGAGFELANQ 242

Qy        252 VIGNVTKAISSYVLAQRVAAGLSTTGAVAALITSSIMLAISPLAFMNAADKFNHANALDE 311
              |:||:|||:|||:||||||||||:|| ||||| |:: ||||||||   ||||||| :|: 
Db        243 VVGNITKAVSSYILAQRVAAGLSSTGPVAALIASTVSLAISPLAFAGIADKFNHAKSLES 302

Qy        312 FAKQFRKFGYDGDHLLAEYQRGVGTIEASLTTISTALGAVSAGVSAAAVGSAVGTPIALL 371
              :|::|:| ||||||:|||||||||| |||:||:| |:||| |:  |||||| || : :|||||
Db        303 YAERFKKLGYDGDNLLAEYQRGTGTIDASVTAINTALAAIAGGVSAAAAGSVIASPIALL 362

Qy        372 VAGVTGLISGILEASKQAMFESVANRLQGKILEWEKQNGGQNYFDKGYDSRYAAYLANNL 431
              |:|:||:|| ||: ||||||| |||::  ||:|||| | |:|||: |||:|| | | :|:
Db        363 VSGITGVISTILQYSKQAMFEHVANKIHNKIVEWEKNNHGKNYFENGYDARYLANLQDNM 422

Qy        432 KFLSELNKELEAERVIAITQQRWDNNIGELAGITKLGERIKSGKAYADAFEDGKKVEAGS 491
              |||  |||||:|||||||||||:||||||:||||::|||:: ||||| ||||:|| ::|  
Db        423 KFLLNLNKELQAERVIAITQQQWDNNIGDLAGISRLGEKVLSGKAYVDAFEEGKHIKADK 482

Qy        492 NITLDAKTGIIDISNSNGKKTQALHFTSPLLTAGTESRERLTNGKYSYINKLKFGRVKNW 551
               : ||:  ||||:|||   ||| : | :||||| ||| |||:  ||| || ||   || :|
Db        483 LVQLDSANGIIDVSNSGKAKTQHILFRTPLLTPGTEHRERVQTGKYEYITKLNINRVDSW 542

Qy        552 QVTDGEASSKLDFSKVIQRV-------AETEGTDEIGLIVNAKAGNDDIFVGQGKMNIDG 604
              ::||| |||  | : |:||:                | |  :|     |:|::||| |   |||
Db        543 KITDGAASSTFDLTNVVQRIGIELDNAGNVTKTKETKIIAKLGEGDDNVFVGSGTTEIDG 602

Qy        605 GDGHDRVFYSKDGGFGNITVDGTSATEAGSYTVNRKVARGDIYHEVVKRQETKVGKRTET 664
              |:|:||| ||: | :| :|:| |  || ||||||| |  |   |||       || | |  
Db        603 GEGYDRVHYSR-GNYGALTIDATKETEQGSYTVNRFVETGKALHEVTSTHTALVGNREEK 661

Qy        665 IQYRDYELRKVGYGYQSTDNLKSVEEVIGSQFNDVFKGSKFNDIFHSGEGDDLLDGGAGD 724
              |:||  :  :   || :  | ||:|||:||:  ||:||||||||| |: |:| | :||  |:
Db        662 IEYR-HSNNQHHAGYYTKDTLKAVEEIIGTSHNDIFKGSKFNDAFNGGDGVDTIDGNDGN 720

Qy        725 DRLFGGKGNDRLSGDEGDDLLDGGSGDDVLNGGAGNDVYIFRKGDGNDTLYDGTGNDKLA 784
              ||||||||:| | |  ||| :||| |:|:|:||:|:|:::  |||||||: |  |||||:
Db        721 DRLFGGKGDDILDGGNGDDFIDGGKGNDLLHGGKGDDIFVHRKGDGNDIITDSDGNDKLS 780

Qy        785 FADANISDIMIERTKEGIIVKRNDHSGSINIPRWY----ITSNLQNYQSNKTDHKIEQLI 840
              |:|:|: |:  |: |: :::  |     :  |  |:        : ||:: | | |||::|
Db        781 FSDSNLKDLTFEKVKHNLVI-TNSKKEKVTIQNWFREADFAKEVPNYKATK-DEKIEEII 838

Qy        841 GKDGSYITSDQIDKILQDKKDGTVITSQELKKLADENKSQKLSASDIASSLNKLVGSMAL 900
              |::|  ||| |:| ::        ||  || |: |:    | |  :: :||:||: |::  
Db        839 GQNGERITSKQVDDLI--AKGNGKITQDELSKVVDNYELLKHS-KNVTNSLDKLISSVSA 895

Qy        901 FGTANSVSSNALQPITQPTQGI 922
              | ::|   : :  | :   | :
Db        896 FTSSNDSRNVLVAPTSMLDQSL 917


RESULT 5
US-07-908-253-2
; Sequence 2, Application US/07908253
; Patent No. 5534256
;  GENERAL INFORMATION:
;    APPLICANT:  POTTER, ANDREW A.
;    APPLICANT:  HARLAND, RICHARD J.
;    TITLE OF INVENTION:  HAEMOPHILUS SOMNUS OUTER MEMBRANE
;    TITLE OF INVENTION:  PROTEIN EXTRACT ENRICHED WITH IRON-REGULATED PROTEINS
;    NUMBER OF SEQUENCES:  6
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  ROBERTA L. ROBINS
;      STREET:  635 BRYANT STREET
;      CITY:  PALO ALTO
;      STATE:  CALIFORNIA
;      COUNTRY:  UNITED STATES OF AMERICA
;      ZIP:  94301
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/908,253
;      FILING DATE:  19920702
;      CLASSIFICATION:  420
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  ROBINS, ROBERTA L.
;      REGISTRATION NUMBER:  33,208
;      REFERENCE/DOCKET NUMBER:  9000-0026
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 617-8999
```

```
;      TELEFAX:  (415) 327-3231
;   INFORMATION FOR SEQ ID NO:  2:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  926 amino acids
;       TYPE:  AMINO ACID
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
US-07-908-253-2

  Query Match             50.1%;  Score 2326.5;  DB 1;  Length 926;
  Best Local Similarity   50.0%;  Pred. No. 1.5e-173;
  Matches  461;  Conservative  174;  Mismatches  262;  Indels   25;  Gaps   12;

Qy         15 LNSTKSGLKNLYLAIPKD--YDPQKGGTLNDFIKAADELGIARLAEEPNHTETAKKSVDT 72
Db          7 LSFPKTGAKKIILYIPQNYQYDTEQGNGLQDLVKAAEELGIEVQREERNNIATAQTSLGT 66

Qy         73 VNQFLSLTQTGIAISATKLEKFLQKHSTNKLAKGLDSVENIDRKLGKASNVLSTLSSFLG 132
Db         67 IQTAIGLTERGIVLSAPQIDKLLQK---TKAGQALGSAESIVQNANKAKTVLSGIQSILG 123

Qy        133 TALAGIELDSLIKKGDAAPDALAKASIDLINEIIGNLSQSTQTIEAFSSQLAKLGSTISQ 192
Db        124 SVLAGMDLDEAL-QNNSNQHALAKAGLELTNSLIENIANSVKTLDEFGEQISQFGSKLQN 182

Qy        193 AKGFSNIGNKLQNL-NFSKTNLGLEIITGLLSGISAGFALADKNASTGKKVAAGFELSNQ 251
Db        183 IKGLGTLGDKLKNIGGLDKAGLGLDVISGLLSGATAALVLADKNASTAKKVGAGFELANQ 242

Qy        252 VIGNVTKAISSYVLAQRVAAGLSTTGAVAALITSSIMLAISPLAFMNAADKFNHANALDE 311
Db        243 VVGNITKAVSSYILAQRVAAGLSSTGPVAALIASTVSLAISPLAFAGIADKFNHAKSLES 302

Qy        312 FAKQFRKFGYDGDHLLAEYQRGVGTIEASLTTISTALGAVSAGVSAAAVGSAVGTPIALL 371
Db        303 YAERFKKLGYDGDNLLAEYQRGTGTIDASVTAINTALAAIAGGVSAAAAGSVIASPIALL 362

Qy        372 VAGVTGLISGILEASKQAMFESVANRLQGKILEWEKQNGGQNYFDKGYDSRYAAYLANNL 431
Db        363 VSGITGVISTILQYSKQAMFEHVANKIHNKIVEWEKNNHGKNYFENGYDARYLANLQDNM 422

Qy        432 KFLSELNKELEAERVIAITQQRWDNNIGELAGITKLGERIKSGKAYADAFEDGKKVEAGS 491
Db        423 KFLLNLNKELQAERVIAITQQQWDNNIGDLAGISRLGEKVLSGKAYVDAFEEGKHIKADK 482

Qy        492 NITLDAKTGIIDISNSNGKKTQALHFTSPLLTAGTESRERLTNGKYSYINKLKFGRVKNW 551
Db        483 LVQLDSANGIIDVSNSGKAKTQHILFRTPLLTPGTEHRERVQTGKYEYITKLNINRVDSW 542

Qy        552 QVTDGEASSKLDFSKVIQRV-------AETEGTDEIGLIVNAKAGNDDIFVGQGKMNIDG 604
Db        543 KITDGAASSTFDLTNVVQRIGIELDNAGNVTKTKETKIIAKLGEGDDNVFVGSGTTEIDG 602

Qy        605 GDGHDRVFYSKDGGFGNITVDGTSATEAGSYTVNRKVARGDIYHEVVKRQETKVGKRTET 664
Db        603 GEGYDRVHYSR-GNYGALTIDATKETEQGSYTVNRFVETGKALHEVTSTHTALVGNREEK 661

Qy        665 IQYRDYELRKVGYGYQSTDNLKSVEEVIGSQFNDVFKGSKFNDIFHSGEGDDLLDGGAGD 724
Db        662 IEYR-HSNNQHHAGYYTKDTLKAVEEIIGTSHNDIFKGSKFNDAFNGGDGVDTIDGNDGN 720

Qy        725 DRLFGGKGNDRLSGDEGDDLLDGGSGDDVLNGGAGNDVYIFRKGDGNDTLYDGTGNDKLA 784
Db        721 DRLFGGKGDDILDGGNGDDFIDGGKGNDLLHGGKGDDIFVHRKGDGNDIITDSDGNDKLS 780

Qy        785 FADANISDIMIERTKEGIIVKRNDHSGSINIPRWY----ITSNLQNYQSNKTDHKIEQLI 840
Db        781 FSDSNLKDLTFEKVKHNLVI-TNSKKEKVTIQNWFREADFAKEVPNYKATK-DEKIEEII 838

Qy        841 GKDGSYITSDQIDKILQDKKDGTVITSQELKKLADENKSQKLSASDIASSLNKLVGSMAL 900
Db        839 GQNGERITSKQVDDLI--AKGNGKITQDELSKVVDNYELLKHS-KNVTNSLDKLISSVSA 895

Qy        901 FGTANSVSSNALQPITQPTQGI 922
Db        896 FTSSNDSRNVLVAPTSMLDQSL 917

RESULT 6
US-08-455-970A-2
; Sequence 2, Application US/08455970A
; Patent No. 5708155
;  GENERAL INFORMATION:
;    APPLICANT:  POTTER, ANDREW A.
;    APPLICANT:  REDMOND, MARK J.
;    APPLICANT:  HUGHES, HUW P.A.
;    TITLE OF INVENTION:  ENHANCED IMMUNOGENICITY USING LEUKOTOXIN
;    TITLE OF INVENTION:  CHIMERAS
;    NUMBER OF SEQUENCES:  15
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  REED & ROBINS
;      STREET:  285 HAMILTON AVENUE, SUITE 200
;      CITY:  PALO ALTO
;      STATE:  CALIFORNIA
;      COUNTRY:  UNITED STATES OF AMERICA
;      ZIP:  94301
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/455,970A
;      FILING DATE:  31-MAY-1995
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/960,932
;      FILING DATE:  14-OCT-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  ROBINS, ROBERTA L.
;      REGISTRATION NUMBER:  33,208
;      REFERENCE/DOCKET NUMBER:  9001-0016.10
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 327-3400
;      TELEFAX:  (415) 327-3231
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  926 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-455-970A-2

  Query Match             50.1%;  Score 2326.5;  DB 1;  Length 926;
  Best Local Similarity   50.0%;  Pred. No. 1.5e-173;
  Matches  461;  Conservative  174;  Mismatches  262;  Indels   25;  Gaps   12;

Qy         15 LNSTKSGLKNLYLAIPKD--YDPQKGGTLNDFIKAADELGIARLAEEPNHTETAKKSVDT 72
Db          7 LSFPKTGAKKIILYIPQNYQYDTEQGNGLQDLVKAAEELGIEVQREERNNIATAQTSLGT 66

Qy         73 VNQFLSLTQTGIAISATKLEKFLQKHSTNKLAKGLDSVENIDRKLGKASNVLSTLSSFLG 132
Db         67 IQTAIGLTERGIVLSAPQIDKLLQK---TKAGQALGSAESIVQNANKAKTVLSGIQSILG 123

Qy        133 TALAGIELDSLIKKGDAAPDALAKASIDLINEIIGNLSQSTQTIEAFSSQLAKLGSTISQ 192
Db        124 SVLAGMDLDEAL-QNNSNQHALAKAGLELTNSLIENIANSVKTLDEFGEQISQFGSKLQN 182

Qy        193 AKGFSNIGNKLQNL-NFSKTNLGLEIITGLLSGISAGFALADKNASTGKKVAAGFELSNQ 251
Db        183 IKGLGTLGDKLKNIGGLDKAGLGLDVISGLLSGATAALVLADKNASTAKKVGAGFELANQ 242

Qy        252 VIGNVTKAISSYVLAQRVAAGLSTTGAVAALITSSIMLAISPLAFMNAADKFNHANALDE 311
Db        243 VVGNITKAVSSYILAQRVAAGLSSTGPVAALIASTVSLAISPLAFAGIADKFNHAKSLES 302
```

```
Qy        312 FAKQFRKFGYDGDHLLAEYQRGVGTIEASLTTISTALGAVSAGVSAAAVGSAVGTPIALL 371
              :|::|:| |||||:||||||| |||:||:| |:||| |:: ||||| || : :|||||
Db        303 YAERFKKLGYDGDNLLAEYQRGTGTIDASVTAINTALAAIAGGVSAAAAGSVIASPIALL 362

Qy        372 VAGVTGLISGILEASKQAMFESVANRLQGKILEWEKQNGGQNYFDKGYDSRYAAYLANNL 431
              |:|:||:|| ||: |||||| |||::  ||:|||| | |:|||: |||:|| | | :|:
Db        363 VSGITGVISTILQYSKQAMFEHVANKIHNKIVEWEKNNHGKNYFENGYDARYLANLQDNM 422

Qy        432 KFLSELNKELEAERVIAITQQRWDNNIGELAGITKLGERIKSGKAYADAFEDGKKVEAGS 491
              |||  |||||:||||||||||:||||||:||||::|||:: ||||| ||||:|| ::|
Db        423 KFLLNLNKELQAERVIAITQQQWDNNIGDLAGISRLGEKVLSGKAYVDAFEEGKHIKADK 482

Qy        492 NITLDAKTGIIDISNSNGKKTQALHFTSPLLTAGTESRERLTNGKYSYINKLKFGRVKNW 551
               : ||:  ||||:|||  ||| : | :|||| ||| |||:  ||| || ||  || :|
Db        483 LVQLDSANGIIDVSNSGKAKTQHILFRTPLLTPGTEHRERVQTGKYEYITKLNINRVDSW 542

Qy        552 QVTDGEASSKLDFSKVIQRV-------AETEGTDEIGLIVNAKAGNDDIFVGQGKMNIDG 604
              ::||| |||  | : |:||:              | |  :|    |:|::||| |   |||
Db        543 KITDGAASSTFDLTNVVQRIGIELDNAGNVTKTKETKIIAKLGEGDDNVFVGSGTTEIDG 602

Qy        605 GDGHDRVFYSKDGGFGNITVDGTSATEAGSYTVNRKVARGDIYHEVVKRQETKVGKRTET 664
              |:|:||| ||: | :| :|:| |  || ||||||| |   |    |||        || | |
Db        603 GEGYDRVHYSR-GNYGALTIDATKETEQGSYTVNRFVETGKALHEVTSTHTALVGNREEK 661

Qy        665 IQYRDYELRKVGYGYQSTDNLKSVEEVIGSQFNDVFKGSKFNDIFHSGEGDDLLDGGAGD 724
              |:|| :   :   || : | ||:|||:||:  ||:||||||| |: |:| | :|| |:
Db        662 IEYR-HSNNQHHAGYYTKDTLKAVEEIIGTSHNDIFKGSKFNDAFNGGDGVDTIDGNDGN 720

Qy        725 DRLFGGKGNDRLSGDEGDDLLDGGSGDDVLNGGAGNDVYIFRKGDGNDTLYDGTGNDKLA 784
              ||||||||:| | |  ||| :||| |:|:|:|| |:|::: ||||||| : | ||||:
Db        721 DRLFGGKGDDILDGGNGDDFIDGGKGNDLLHGGKGDDIFVHRKGDGNDIITDSDGNDKLS 780

Qy        785 FADANISDIMIERTKEGIIVKRNDHSGSINIPRWY----ITSNLQNYQSNKTDHKIEQLI 840
              |:|:|: |:  |:| |  ::: |      : |  |:           : ||:: | | |||::|
Db        781 FSDSNLKDLTFEKVKHNLVI-TNSKKEKVTIQNWFREADFAKEVPNYKATK-DEKIEEII 838

Qy        841 GKDGSYITSDQIDKILQDKKDGTVITSQELKKLADENKSQKLSASDIASSLNKLVGSMAL 900
              |::| ||| |:| ::    |    ||  ||  |:  |   :   | | :: :|:||:| |::
Db        839 GQNGERITSKQVDDLI--AKGNGKITQDELSKVVDNYELLKHS-KNVTNSLDKLISSVSA 895

Qy        901 FGTANSVSSNALQPITQPTQGI 922
              | ::|   :  : | :   | :
Db        896 FTSSNDSRNVLVAPTSMLDQSL 917
```

RESULT 7
US-08-387-156-6
; Sequence 6, Application US/08387156
; Patent No. 5723129
; GENERAL INFORMATION:
;   APPLICANT: POTTER, ANDREW A.
;   APPLICANT: REDMOND, MARK J.
;   APPLICANT: HUGHES, HUW P.A.
;   TITLE OF INVENTION: GnRH-LEUKOTOXIN CHIMERAS
;   NUMBER OF SEQUENCES: 28
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: REED & ROBINS
;   STREET: 635 BRYANT STREET
;   CITY: PALO ALTO
;   STATE: CALIFORNIA
;   COUNTRY: UNITED STATES OF AMERICA
;   ZIP: 94301
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.25
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/387,156
;   FILING DATE: 10-FEB-1995
;   CLASSIFICATION: 424
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US 07/960,932
;   FILING DATE: 14-OCT-1992
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US 07/779,171
;   FILING DATE: 16-OCT-1991
;   ATTORNEY/AGENT INFORMATION:
;   NAME: ROBINS, ROBERTA L.
;   REGISTRATION NUMBER: 33,208
;   REFERENCE/DOCKET NUMBER: 9001-0016.21
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (415) 617-8999
;   TELEFAX: (415) 327-3231
; INFORMATION FOR SEQ ID NO: 6:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 926 amino acids
;   TYPE: amino acid
;   TOPOLOGY: linear
;   MOLECULE TYPE: protein
US-08-387-156-6

Query Match 50.1%; Score 2326.5; DB 1; Length 926;
Best Local Similarity 50.0%; Pred. No. 1.5e-173;
Matches 461; Conservative 174; Mismatches 262; Indels 25; Gaps 12;

```
Qy         15 LNSTKSGLKNLYLAIPKD--YDPQKGGTLNDFIKAADELGIARLAEEPNHTETAKKSVDT 72
              |:  |:| | : |  ||::  |   | |  :|||:|||       ||  |:  |:  |
Db          7 LSFPKTGAKKIILYIPQNYQYDTEQGNGLQDLVKAAEELGIEVQREERNNIATAQTSLGT 66

Qy         73 VNQFLSLTQTGIAISATKLEKFLQKHSTNKLAKGLDSVENIDRKLGKASNVLSTLSSFLG 132
               :   : ||: || :||  :::| |||       |   | | | |:|  :     ||  |||  :  | ||
Db         67 IQTAIGLTERGIVLSAPQIDKLLQK---TKAGQALGSAESIVQNANKAKTVLSGIQSILG 123

Qy        133 TALAGIELDSLIKKGDAAPDALAKASIDLINEIIGNLSQSTQTIEAFSSQLAKLGSTISQ 192
              : |||::||  : : ::   ||||| ::| | :| |:: | :|:: |  |::: || :
Db        124 SVLAGMDLDEAL-QNNSNQHALAKAGLELTNSLIENIANSVKTLDEFGEQISQFGSKLQN 182

Qy        193 AKGFSNIGNKLQNL-NFSKTNLGLEIITGLLSGISAGFALADKNASTGKKVAAGFELSNQ 251
               ||   :|:||:|:    |   |||::|:||||| :|  ||||||| ||| ||||:||
Db        183 IKGLGTLGDKLKNIGGLDKAGLGLDVISGLLSGATAALVLADKNASTAKKVGAGFELANQ 242

Qy        252 VIGNVTKAISSYVLAQRVAAGLSTTGAVAALITSSIMLAISPLAFMNAADKFNHANALDE 311
              |:||:||:|||:||||||||||:|||  ||||  |:  |||||||:   ||||||| :|:
Db        243 VVGNITKAVSSYILAQRVAAGLSSTGPVAALIASTVSLAISPLAFAGIADKFNHAKSLES 302

Qy        312 FAKQFRKFGYDGDHLLAEYQRGVGTIEASLTTISTALGAVSAGVSAAAVGSAVGTPIALL 371
              :|::|:| |||||:||||||| |||:||:| |:||| |:: ||||| || : :|||||
Db        303 YAERFKKLGYDGDNLLAEYQRGTGTIDASVTAINTALAAIAGGVSAAAAGSVIASPIALL 362

Qy        372 VAGVTGLISGILEASKQAMFESVANRLQGKILEWEKQNGGQNYFDKGYDSRYAAYLANNL 431
              |:|:||:|| ||: |||||| |||::  ||:|||| | |:|||: |||:|| | | :|:
Db        363 VSGITGVISTILQYSKQAMFEHVANKIHNKIVEWEKNNHGKNYFENGYDARYLANLQDNM 422

Qy        432 KFLSELNKELEAERVIAITQQRWDNNIGELAGITKLGERIKSGKAYADAFEDGKKVEAGS 491
              |||  |||||:||||||||||:||||||:||||::|||:: ||||| ||||:|| ::|
Db        423 KFLLNLNKELQAERVIAITQQQWDNNIGDLAGISRLGEKVLSGKAYVDAFEEGKHIKADK 482

Qy        492 NITLDAKTGIIDISNSNGKKTQALHFTSPLLTAGTESRERLTNGKYSYINKLKFGRVKNW 551
               : ||:  ||||:|||  ||| : | :|||| ||| |||:  ||| || ||  || :|
Db        483 LVQLDSANGIIDVSNSGKAKTQHILFRTPLLTPGTEHRERVQTGKYEYITKLNINRVDSW 542

Qy        552 QVTDGEASSKLDFSKVIQRV-------AETEGTDEIGLIVNAKAGNDDIFVGQGKMNIDG 604
              ::||| |||  | : |:||:              | |  :|    |:|::||| |   |||
Db        543 KITDGAASSTFDLTNVVQRIGIELDNAGNVTKTKETKIIAKLGEGDDNVFVGSGTTEIDG 602

Qy        605 GDGHDRVFYSKDGGFGNITVDGTSATEAGSYTVNRKVARGDIYHEVVKRQETKVGKRTET 664
              |:|:||| ||: | :| :|:| |  || ||||||| |   |    |||        || | |
Db        603 GEGYDRVHYSR-GNYGALTIDATKETEQGSYTVNRFVETGKALHEVTSTHTALVGNREEK 661

Qy        665 IQYRDYELRKVGYGYQSTDNLKSVEEVIGSQFNDVFKGSKFNDIFHSGEGDDLLDGGAGD 724
              |:|| :   :   || : | ||:|||:||:  ||:||||||| |: |:| | :|| |:
Db        662 IEYR-HSNNQHHAGYYTKDTLKAVEEIIGTSHNDIFKGSKFNDAFNGGDGVDTIDGNDGN 720

Qy        725 DRLFGGKGNDRLSGDEGDDLLDGGSGDDVLNGGAGNDVYIFRKGDGNDTLYDGTGNDKLA 784
```

```
Db        721 DRLFGGKGDDILDGGNGDDFIDGGKGNDLLHGGKGDDIFVHRKGDGNDIITDSDGNDKLS 780

Qy        785 FADANISDIMIERTKEGIIVKRNDHSGSINIPRWY----ITSNLQNYQSNKTDHKIEQLI 840

Db        781 FSDSNLKDLTFEKVKHNLVI-TNSKKEKVTIQNWFREADFAKEVPNYKATK-DEKIEEII 838

Qy        841 GKDGSYITSDQIDKILQDKKDGTVITSQELKKLADENKSQKLSASDIASSLNKLVGSMAL 900

Db        839 GQNGERITSKQVDDLI--AKGNGKITQDELSKVVDNYELLKHS-KNVTNSLDKLISSVSA 895

Qy        901 FGTANSVSSNALQPITQPTQGI 922

Db        896 FTSSNDSRNVLVAPTSMLDQSL 917
```

RESULT 8
US-08-694-865-6
; Sequence 6, Application US/08694865
; Patent No. 5837268
; GENERAL INFORMATION:
; APPLICANT: POTTER, ANDREW A.
; APPLICANT: MANNS, JOHN G.
; TITLE OF INVENTION: GnRH-LEUKOTOXIN CHIMERAS
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: REED & ROBINS LLP
; STREET: 285 HAMILTON AVENUE, SUITE 200
; CITY: PALO ALTO
; STATE: CA
; COUNTRY: USA
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/694,865
; FILING DATE: 09-AUG-1996
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: MCCRACKEN, THOMAS P.
; REGISTRATION NUMBER: 38,548
; REFERENCE/DOCKET NUMBER: 9001-0016.22
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415)327-3400
; TELEFAX: (415)327-3231
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 926 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-694-865-6

```
  Query Match             50.1%;  Score 2326.5;  DB 2;  Length 926;
  Best Local Similarity   50.0%;  Pred. No. 1.5e-173;
  Matches  461;  Conservative  174;  Mismatches  262;  Indels   25;  Gaps   12;

Qy         15 LNSTKSGLKNLYLAIPKD--YDPQKGGTLNDFIKAADELGIARLAEEPNHTETAKKSVDT 72

Db          7 LSFPKTGAKKIILYIPQNYQYDTEQGNGLQDLVKAAEELGIEVQREERNNIATAQTSLGT 66

Qy         73 VNQFLSLTQTGIAISATKLEKFLQKHSTNKLAKGLDSVENIDRKLGKASNVLSTLSSFLG 132

Db         67 IQTAIGLTERGIVLSAPQIDKLLQK---TKAGQALGSAESIVQNANKAKTVLSGIQSILG 123

Qy        133 TALAGIELDSLIKKGDAAPDALAKASIDLINEIIGNLSQSTQTIEAFSSQLAKLGSTISQ 192

Db        124 SVLAGMDLDEAL-QNNSNQHALAKAGLELTNSLIENIANSVKTLDEFGEQISQFGSKLQN 182

Qy        193 AKGFSNIGNKLQNL-NFSKTNLGLEIITGLLSGISAGFALADKNASTGKKVAAGFELSNQ 251

Db        183 IKGLGTLGDKLKNIGGLDKAGLGLDVISGLLSGATAALVLADKNASTAKKVGAGFELANQ 242

Qy        252 VIGNVTKAISSYVLAQRVAAGLSTTGAVAALITSSIMLAISPLAFMNAADKFNHANALDE 311

Db        243 VVGNITKAVSSYILAQRVAAGLSSTGPVAALIASTVSLAISPLAFAGIADKFNHAKSLES 302

Qy        312 FAKQFRKFGYDGDHLLAEYQRGVGTIEASLTTISTALGAVSAGVSAAAVGSAVGTPIALL 371

Db        303 YAERFKKLGYDGDNLLAEYQRGTGTIDASVTAINTALAAIAGGVSAAAAGSVIASPIALL 362

Qy        372 VAGVTGLISGILEASKQAMFESVANRLQGKILEWEKQNGGQNYFDKGYDSRYAAYLANNL 431

Db        363 VSGITGVISTILQYSKQAMFEHVANKIHNKIVEWEKNNHGKNYFENGYDARYLANLQDNM 422

Qy        432 KFLSELNKELEAERVIAITQQRWDNNIGELAGITKLGERIKSGKAYADAFEDGKKVEAGS 491

Db        423 KFLLNLNKELQAERVIAITQQQWDNNIGDLAGISRLGEKVLSGKAYVDAFEEGKHIKADK 482

Qy        492 NITLDAKTGIIDISNSNGKKTQALHFTSPLLTAGTESRERLTNGKYSYINKLKFGRVKNW 551

Db        483 LVQLDSANGIIDVSNSGKAKTQHILFRTPLLTPGTEHRERVQTGKYEYITKLNINRVDSW 542

Qy        552 QVTDGEASSKLDFSKVIQRV-------AETEGTDEIGLIVNAKAGNDDIFVGQGKMNIDG 604

Db        543 KITDGAASSTFDLTNVVQRIGIELDNAGNVTKTKETKIIAKLGEGDDNVFVGSGTTEIDG 602

Qy        605 GDGHDRVFYSKDGGFGNITVDGTSATEAGSYTVNRKVARGDIYHEVVKRQETKVGKRTET 664

Db        603 GEGYDRVHYSR-GNYGALTIDATKETEQGSYTVNRFVETGKALHEVTSTHTALVGNREEK 661

Qy        665 IQYRDYELRKVGYGYQSTDNLKSVEEVIGSQFNDVFKGSKFNDIFHSGEGDDLLDGGAGD 724

Db        662 IEYR-HSNNQHHAGYYTKDTLKAVEEIIGTSHNDIFKGSKFNDAFNGGDGVDTIDGNDGN 720

Qy        725 DRLFGGKGNDRLSGDEGDDLLDGGSGDDVLNGGAGNDVYIFRKGDGNDTLYDGTGNDKLA 784

Db        721 DRLFGGKGDDILDGGNGDDFIDGGKGNDLLHGGKGDDIFVHRKGDGNDIITDSDGNDKLS 780

Qy        785 FADANISDIMIERTKEGIIVKRNDHSGSINIPRWY----ITSNLQNYQSNKTDHKIEQLI 840

Db        781 FSDSNLKDLTFEKVKHNLVI-TNSKKEKVTIQNWFREADFAKEVPNYKATK-DEKIEEII 838

Qy        841 GKDGSYITSDQIDKILQDKKDGTVITSQELKKLADENKSQKLSASDIASSLNKLVGSMAL 900

Db        839 GQNGERITSKQVDDLI--AKGNGKITQDELSKVVDNYELLKHS-KNVTNSLDKLISSVSA 895

Qy        901 FGTANSVSSNALQPITQPTQGI 922

Db        896 FTSSNDSRNVLVAPTSMLDQSL 917
```

RESULT 9
US-08-878-748-6
; Sequence 6, Application US/08878748
; Patent No. 5969126
; GENERAL INFORMATION:
; APPLICANT: POTTER, ANDREW A.
; APPLICANT: REDMOND, MARK J.
; APPLICANT: HUGHES, HUW P.A.
; TITLE OF INVENTION: GnRH-LEUKOTOXIN CHIMERAS
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: REED & ROBINS
; STREET: 635 BRYANT STREET
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

```
;     SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/878,748
;      FILING DATE:  19-JUN-1997
;      CLASSIFICATION:  536
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/387,156
;      FILING DATE:  10-FEB-1995
;      APPLICATION NUMBER:  US 07/960,932
;      FILING DATE:  14-OCT-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/779,171
;      FILING DATE:  16-OCT-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  ROBINS, ROBERTA L.
;      REGISTRATION NUMBER:  33,208
;      REFERENCE/DOCKET NUMBER:  9001-0016.21
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 617-8999
;      TELEFAX:  (415) 327-3231
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  926 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-878-748-6

  Query Match             50.1%;  Score 2326.5;  DB 2;  Length 926;
  Best Local Similarity   50.0%;  Pred. No. 1.5e-173;
  Matches  461;  Conservative  174;  Mismatches  262;  Indels   25;  Gaps   12;

Qy         15 LNSTKSGLKNLYLAIPKD--YDPQKGGTLNDFIKAADELGIARLAEEPNHTETAKKSVDT 72
Db          7 LSFPKTGAKKIILYIPQNYQYDTEQGNGLQDLVKAAEELGIEVQREERNNIATAQTSLGT 66

Qy         73 VNQFLSLTQTGIAISATKLEKFLQKHSTNKLAKGLDSVENIDRKLGKASNVLSTLSSFLG 132
Db         67 IQTAIGLTERGIVLSAPQIDKLLQK---TKAGQALGSAESIVQNANKAKTVLSGIQSILG 123

Qy        133 TALAGIELDSLIKKGDAAPDALAKASIDLINEIIGNLSQSTQTIEAFSSQLAKLGSTISQ 192
Db        124 SVLAGMDLDEAL-QNNSNQHALAKAGLELTNSLIENIANSVKTLDEFGEQISQFGSKLQN 182

Qy        193 AKGFSNIGNKLQNL-NFSKTNLGLEIITGLLSGISAGFALADKNASTGKKVAAGFELSNQ 251
Db        183 IKGLGTLGDKLKNIGGLDKAGLGLDVISGLLSGATAALVLADKNASTAKKVGAGFELANQ 242

Qy        252 VIGNVTKAISSYVLAQRVAAGLSTTGAVAALITSSIMLAISPLAFMNAADKFNHANALDE 311
Db        243 VVGNITKAVSSYILAQRVAAGLSSTGPVAALIASTVSLAISPLAFAGIADKFNHAKSLES 302

Qy        312 FAKQFRKFGYDGDHLLAEYQRGVGTIEASLTTISTALGAVSAGVSAAAVGSAVGTPIALL 371
Db        303 YAERFKKLGYDGDNLLAEYQRGTGTIDASVTAINTALAAIAGGVSAAAAGSVIASPIALL 362

Qy        372 VAGVTGLISGILEASKQAMFESVANRLQGKILEWEKQNGGQNYFDKGYDSRYAAYLANNL 431
Db        363 VSGITGVISTILQYSKQAMFEHVANKIHNKIVEWEKNNHGKNYFENGYDARYLANLQDNM 422

Qy        432 KFLSELNKELEAERVIAITQQRWDNNIGELAGITKLGERIKSGKAYADAFEDGKKVEAGS 491
Db        423 KFLLNLNKELQAERVIAITQQQWDNNIGDLAGISRLGEKVLSGKAYVDAFEEGKHIKADK 482

Qy        492 NITLDAKTGIIDISNSNGKKTQALHFTSPLLTAGTESRERLTNGKKYSYINKLKFGRVKNW 551
Db        483 LVQLDSANGIIDVSNSGKAKTQHILFRTPLLTPGTEHRERVQTGKYEYITKLNINRVDSW 542

Qy        552 QVTDGEASSKLDFSKVIQRV-------AETEGTDEIGLIVNAKAGNDDIFVGQGKMNIDG 604
Db        543 KITDGAASSTFDLTNVVQRIGIELDNAGNVTKTKETKIIAKLGEGDDNVFVGSGTTEIDG 602

Qy        605 GDGHDRVFYSKDGGFGNITVDGTSATEAGSYTVNRKVARGDIYHEVVKRQETKVGKRTET 664

Db        603 GEGYDRVHYSR-GNYGALTIDATKETEQGSYTVNRFVETGKALHEVTSTHTALVGNREEK 661

Qy        665 IQYRDYELRKVGYGYQSTDNLKSVEEVIGSQFNDVFKGSKFNDIFHSGEGDDLLDGGAGD 724
Db        662 IEYR-HSNNQHHAGYYTKDTLKAVEEIIGTSHNDIFKGSKFNDAFNGGDGVDTIDGNDGN 720

Qy        725 DRLFGGKGNDRLSGDEGDDLLDGGSGDDVLNGGAGNDVYIFRKGDGNDTLYDGTGNDKLA 784
Db        721 DRLFGGKGDDILDGGNGDDFIDGGKGNDLLHGGKGDDIFVHRKGDGNDIITDSDGNDKLS 780

Qy        785 FADANISDIMIERTKEGIIVKRNDHSGSINIPRWY----ITSNLQNYQSNKTDHKIEQLI 840
Db        781 FSDSNLKDLTFEKVKHNLVI-TNSKKEKVTIQNWFREADFAKEVPNYKATK-DEKIEEII 838

Qy        841 GKDGSYITSDQIDKILQDKKDGTVITSQELKKLADENKSQKLSASDIASSLNKLVGSMAL 900
Db        839 GQNGERITSKQVDDLI--AKGNGKITQDELSKVVDNYELLKHS-KNVTNSLDKLISSVSA 895

Qy        901 FGTANSVSSNALQPITQPTQGI 922
Db        896 FTSSNDSRNVLVAPTSMLDQSL 917

RESULT 10
US-08-535-837-2
; Sequence 2, Application US/08535837
; Patent No. 5985289
;  GENERAL INFORMATION:
;    APPLICANT:  POTTER, ANDREW A.
;    APPLICANT:  HARLAND, RICHARD J.
;    TITLE OF INVENTION:  HAEMOPHILUS SOMNUS OUTER MEMBRANE
;    TITLE OF INVENTION:  PROTEIN EXTRACT ENRICHED WITH IRON-REGULATED PROTEINS
;    NUMBER OF SEQUENCES:  6
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  ROBERTA L. ROBINS
;      STREET:  285 HAMILTON AVENUE, SUITE 200
;      CITY:  PALO ALTO
;      STATE:  CALIFORNIA
;      COUNTRY:  UNITED STATES OF AMERICA
;      ZIP:  94301
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/535,837
;      FILING DATE:  27-SEP-1995
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  ROBINS, ROBERTA L.
;      REGISTRATION NUMBER:  33,208
;      REFERENCE/DOCKET NUMBER:  9001-0026.01
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 327-3400
;      TELEFAX:  (415) 327-3231
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  926 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-535-837-2

  Query Match             50.1%;  Score 2326.5;  DB 2;  Length 926;
  Best Local Similarity   50.0%;  Pred. No. 1.5e-173;
  Matches  461;  Conservative  174;  Mismatches  262;  Indels   25;  Gaps   12;

Qy         15 LNSTKSGLKNLYLAIPKD--YDPQKGGTLNDFIKAADELGIARLAEEPNHTETAKKSVDT 72
Db          7 LSFPKTGAKKIILYIPQNYQYDTEQGNGLQDLVKAAEELGIEVQREERNNIATAQTSLGT 66
```

```
Qy      73 VNQFLSLTQTGIAISATKLEKFLQKHSTNKLAKGLDSVENIDRKLGKASNVLSTLSSFLG 132
Db      67 IQTAIGLTERGIVLSAPQIDKLLQK---TKAGQALGSAESIVQNANKAKTVLSGIQSILG 123

Qy     133 TALAGIELDSLIKKGDAAPDALAKASIDLINEIIGNLSQSTQTIEAFSSQLAKLGSTISQ 192
Db     124 SVLAGMDLDEAL-QNNSNQHALAKAGLELTNSLIENIANSVKTLDEFGEQISQFGSKLQN 182

Qy     193 AKGFSNIGNKLQNL-NFSKTNLGLEIITGLLSGISAGFALADKNASTGKKVAAGFELSNQ 251
Db     183 IKGLGTLGDKLKNIGGLDKAGLGLDVISGLLSGATAALVLADKNASTAKKVGAGFELANQ 242

Qy     252 VIGNVTKAISSYVLAQRVAAGLSTTGAVAALITSSIMLAISPLAFMNAADKFNHANALDE 311
Db     243 VVGNITKAVSSYILAQRVAAGLSSTGPVAALIASTVSLAISPLAFAGIADKFNHAKSLES 302

Qy     312 FAKQFRKFGYDGDHLLAEYQRGVGTIEASLTTISTALGAVSAGVSAAAVGSAVGTPIALL 371
Db     303 YAERFKKLGYDGDNLLAEYQRGTGTIDASVTAINTALAAIAGGVSAAAAGSVIASPIALL 362

Qy     372 VAGVTGLISGILEASKQAMFESVANRLQGKILEWEKQNGGQNYFDKGYDSRYAAYLANNL 431
Db     363 VSGITGVISTILQYSKQAMFEHVANKIHNKIVEWEKNNHGKNYFENGYDARYLANLQDNM 422

Qy     432 KFLSELNKELEAERVIAITQQRWDNNIGELAGITKLGERIKSGKAYADAFEDGKKVEAGS 491
Db     423 KFLLNLNKELQAERVIAITQQQWDNNIGDLAGISRLGEKVLSGKAYVDAFEEGKHIKADK 482

Qy     492 NITLDAKTGIIDISNSNGKKTQALHFTSPLLTAGTESRERLTNGKYSYINKLKFGRVKNW 551
Db     483 LVQLDSANGIIDVSNSGKAKTQHILFRTPLLTPGTEHRERVQTGKYEYITKLNINRVDSW 542

Qy     552 QVTDGEASSKLDFSKVIQRV-------AETEGTDEIGLIVNAKAGNDDIFVGQGKMNIDG 604
Db     543 KITDGAASSTFDLTNVVQRIGIELDNAGNVTKTKETKIIAKLGEGDDNVFVGSGTTEIDG 602

Qy     605 GDGHDRVFYSKDGGFGNITVDGTSATEAGSYTVNRKVARGDIYHEVVKRQETKVGKRTET 664
Db     603 GEGYDRVHYSR-GNYGALTIDATKETEQGSYTVNRFVETGKALHEVTSTHTALVGNREEK 661

Qy     665 IQYRDYELRKVGYGYQSTDNLKSVEEVIGSQFNDVFKGSKFNDIFHSGEGDDLLDGGAGD 724
Db     662 IEYR-HSNNQHHAGYYTKDTLKAVEEIIGTSHNDIFKGSKFNDAFNGGDGVDTIDGNDGN 720

Qy     725 DRLFGGKGNDRLSGDEGDDLLDGGSGDDVLNGGAGNDVYIFRKGDGNDTLYDGTGNDKLA 784
Db     721 DRLFGGKGDDILDGGNGDDFIDGGKGNDLLHGGKGDDIFVHRKGDGNDIITDSDGNDKLS 780

Qy     785 FADANISDIMIERTKEGIIVKRNDHSGSINIPRWY----ITSNLQNYQSNKTDHKIEQLI 840
Db     781 FSDSNLKDLTFEKVKHNLVI-TNSKKEKVTIQNWFREADFAKEVPNYKATK-DEKIEEII 838

Qy     841 GKDGSYITSDQIDKILQDKKDGTVITSQELKKLADENKSQKLSASDIASSLNKLVGSMAL 900
Db     839 GQNGERITSKQVDDLI--AKGNGKITQDELSKVVDNYELLKHS-KNVTNSLDKLISSVSA 895

Qy     901 FGTANSVSSNALQPITQPTQGI 922
Db     896 FTSSNDSRNVLVAPTSMLDQSL 917
```

RESULT 11
US-09-124-491-6
; Sequence 6, Application US/09124491
; Patent No. 6022960
; GENERAL INFORMATION:
; APPLICANT: POTTER, ANDREW A.
; APPLICANT: MANNS, JOHN G.
; TITLE OF INVENTION: GnRH-LEUKOTOXIN CHIMERAS
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: REED & ROBINS LLP
; STREET: 285 HAMILTON AVENUE, SUITE 200
; CITY: PALO ALTO
; STATE: CA
; COUNTRY: USA
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/124,491
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/694,865
; FILING DATE: 09-AUG-1996
; APPLICATION NUMBER: US 08/387,156
; FILING DATE: 10-FEB-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/960,932
; FILING DATE: 14-OCT-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/779,171
; FILING DATE: 16-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: MCCRACKEN, THOMAS P.
; REGISTRATION NUMBER: 38,548
; REFERENCE/DOCKET NUMBER: 9001-0016.22
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415)327-3400
; TELEFAX: (415)327-3231
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 926 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-124-491-6

```
Query Match             50.1%;  Score 2326.5;  DB 3;  Length 926;
Best Local Similarity   50.0%;  Pred. No. 1.5e-173;
Matches 461;  Conservative 174;  Mismatches 262;  Indels 25;  Gaps 12;

Qy      15 LNSTKSGLKNLYLAIPKD--YDPQKGGTLNDFIKAADELGIARLAEEPNHTETAKKSVDT 72
Db       7 LSFPKTGAKKIILYIPQNYQYDTEQGNGLQDLVKAAEELGIEVQREERNNIATAQTSLGT 66

Qy      73 VNQFLSLTQTGIAISATKLEKFLQKHSTNKLAKGLDSVENIDRKLGKASNVLSTLSSFLG 132
Db      67 IQTAIGLTERGIVLSAPQIDKLLQK---TKAGQALGSAESIVQNANKAKTVLSGIQSILG 123

Qy     133 TALAGIELDSLIKKGDAAPDALAKASIDLINEIIGNLSQSTQTIEAFSSQLAKLGSTISQ 192
Db     124 SVLAGMDLDEAL-QNNSNQHALAKAGLELTNSLIENIANSVKTLDEFGEQISQFGSKLQN 182

Qy     193 AKGFSNIGNKLQNL-NFSKTNLGLEIITGLLSGISAGFALADKNASTGKKVAAGFELSNQ 251
Db     183 IKGLGTLGDKLKNIGGLDKAGLGLDVISGLLSGATAALVLADKNASTAKKVGAGFELANQ 242

Qy     252 VIGNVTKAISSYVLAQRVAAGLSTTGAVAALITSSIMLAISPLAFMNAADKFNHANALDE 311
Db     243 VVGNITKAVSSYILAQRVAAGLSSTGPVAALIASTVSLAISPLAFAGIADKFNHAKSLES 302

Qy     312 FAKQFRKFGYDGDHLLAEYQRGVGTIEASLTTISTALGAVSAGVSAAAVGSAVGTPIALL 371
Db     303 YAERFKKLGYDGDNLLAEYQRGTGTIDASVTAINTALAAIAGGVSAAAAGSVIASPIALL 362

Qy     372 VAGVTGLISGILEASKQAMFESVANRLQGKILEWEKQNGGQNYFDKGYDSRYAAYLANNL 431
Db     363 VSGITGVISTILQYSKQAMFEHVANKIHNKIVEWEKNNHGKNYFENGYDARYLANLQDNM 422

Qy     432 KFLSELNKELEAERVIAITQQRWDNNIGELAGITKLGERIKSGKAYADAFEDGKKVEAGS 491
```

```
Db        423 KFLLNLNKELQAERVIAITQQQWDNNIGDLAGISRLGEKVLSGKAYVDAFEEGKHIKADK 482

Qy        492 NITLDAKTGIIDISNSNGKKTQALHFTSPLLTAGTESRERLTNGKYSYINKLKFGRVKNW 551
                : ||:  ||||:|||    |||  : | :||||  ||| |||:   ||| || ||      || :|
Db        483 LVQLDSANGIIDVSNSGKAKTQHILFRTPLLTPGTEHRERVQTGKYEYITKLNINRVDSW 542

Qy        552 QVTDGEASSKLDFSKVIQRV-------AETEGTDEIGLIVNAKAGNDDIFVGQGKMNIDG 604
              ::||| |||   |  : |:||:                 | |   :|       |:|::||| |      |||
Db        543 KITDGAASSTFDLTNVVQRIGIELDNAGNVTKTKETKIIAKLGEGDDNVFVGSGTTEIDG 602

Qy        605 GDGHDRVFYSKDGGFGNITVDGTSATEAGSYTVNRKVARGDIYHEVVKRQETKVGKRTET 664
              |:|:||| ||:  |  :|   :|:|  |    || ||||||| |     |    |||          || |  |
Db        603 GEGYDRVHYSR-GNYGALTIDATKETEQGSYTVNRFVETGKALHEVTSTHTALVGNREEK 661

Qy        665 IQYRDYELRKVGYGYQSTDNLKSVEEVIGSQFNDVFKGSKFNDIFHSGEGDDLLDGGAGD 724
              |:||  :      :      || :  | ||:|||:||:   ||:||||||||  |:  |:| |  :|| |:
Db        662 IEYR-HSNNQHHAGYYTKDTLKAVEEIIGTSHNDIFKGSKFNDAFNGGDGVDTIDGNDGN 720

Qy        725 DRLFGGKGNDRLSGDEGDDLLDGGSGDDVLNGGAGNDVYIFRKGDGNDTLYDGTGNDKLA 784
              |||||||:|  |  |    |||  :|||  |:|:|:|:|| |:|::::  ||||||  :  |   |||||:
Db        721 DRLFGGKGDDILDGGNGDDFIDGGKGNDLLHGGKGDDIFVHRKGDGNDIITDSDGNDKLS 780

Qy        785 FADANISDIMIERTKEGIIVKRNDHSGSINIPRWY----ITSNLQNYQSNKTDHKIEQLI 840
              |:|:|: |:  |: |   :::  |           :  |  |:             : ||::  |  | |||::|
Db        781 FSDSNLKDLTFEKVKHNLVI-TNSKKEKVTIQNWFREADFAKEVPNYKATK-DEKIEEII 838

Qy        841 GKDGSYITSDQIDKILQDKKDGTVITSQELKKLADENKSQKLSASDIASSLNKLVGSMAL 900
              |::|   |||  |:|  ::      |          ||   || |:  |    :   | |     :: :||:||: |::
Db        839 GQNGERITSKQVDDLI--AKGNGKITQDELSKVVDNYELLKHS-KNVTNSLDKLISSVSA 895

Qy        901 FGTANSVSSNALQPITQPTQGI 922
              | ::|    :   :  |  :    | :
Db        896 FTSSNDSRNVLVAPTSMLDQSL 917
```

RESULT 12
US-09-383-912-6
; Sequence 6, Application US/09383912
; Patent No. 6521746
; GENERAL INFORMATION:
; APPLICANT: POTTER, ANDREW A.
; APPLICANT: MANNS, JOHN G.
; TITLE OF INVENTION: GnRH-LEUKOTOXIN CHIMERAS
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: REED & ROBINS LLP
; STREET: 285 HAMILTON AVENUE, SUITE 200
; CITY: PALO ALTO
; STATE: CA
; COUNTRY: USA
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/383,912
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/694,865
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: MCCRACKEN, THOMAS P.
; REGISTRATION NUMBER: 38,548
; REFERENCE/DOCKET NUMBER: 9001-0016.22
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415)327-3400
; TELEFAX: (415)327-3231
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 926 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-383-912-6

Query Match 50.1%; Score 2326.5; DB 4; Length 926;
Best Local Similarity 50.0%; Pred. No. 1.5e-173;
Matches 461; Conservative 174; Mismatches 262; Indels 25; Gaps 12;

```
Qy         15 LNSTKSGLKNLYLAIPKD--YDPQKGGTLNDFIKAADELGIARLAEEPNHTETAKKSVDT 72
              |:  |:| | : | ||::  || ::|  | | :|||:||||       || |:  ||: |: |
Db          7 LSFPKTGAKKIILYIPQNYQYDTEQGNGLQDLVKAAEELGIEVQREERNNIATAQTSLGT 66

Qy         73 VNQFLSLTQTGIAISATKLEKFLQKHSTNKLAKGLDSVENIDRKLGKASNVLSTLSSFLG 132
              :   : ||: || :|| :::| |||        | :  | | |:|  :      ||  |||  : | ||
Db         67 IQTAIGLTERGIVLSAPQIDKLLQK---TKAGQALGSAESIVQNANKAKTVLSGIQSILG 123

Qy        133 TALAGIELDSLIKKGDAAPDALAKASIDLINEIIGNLSQSTQTIEAFSSQLAKLGSTISQ 192
              : |||::||  : : ::    |||||  ::| | :| |:: | :|:: |  |::: || :
Db        124 SVLAGMDLDEAL-QNNSNQHALAKAGLELTNSLIENIANSVKTLDEFGEQISQFGSKLQN 182

Qy        193 AKGFSNIGNKLQNL-NFSKTNLGLEIITGLLSGISAGFALADKNASTGKKVAAGFELSNQ 251
              ||   :|:||:|:   |  |||:|:||||| :  ||||||| ||| |||||:||
Db        183 IKGLGTLGDKLKNIGGLDKAGLGLDVISGLLSGATAALVLADKNASTAKKVGAGFELANQ 242

Qy        252 VIGNVTKAISSYVLAQRVAAGLSTTGAVAALITSSIMLAISPLAFMNAADKFNHANALDE 311
              |:||:|||:|||:||||||||||:|| |||||  |:: ||||||     ||||||  :|:
Db        243 VVGNITKAVSSYILAQRVAAGLSSTGPVAALIASTVSLAISPLAFAGIADKFNHAKSLES 302

Qy        312 FAKQFRKFGYDGDHLLAEYQRGVGTIEASLTTISTALGAVSAGVSAAAVGSAVGTPIALL 371
              :|::|:| ||||:|||||||| |||:||:| |:||| |:: ||||| || : :||||
Db        303 YAERFKKLGYDGDNLLAEYQRGTGTIDASVTAINTALAAIAGGVSAAAAGSVIASPIALL 362

Qy        372 VAGVTGLISGILEASKQAMFESVANRLQGKILEWEKQNGGQNYFDKGYDSRYAAYLANNL 431
              |:|:||:|| ||: ||||||| |||::    |:||||  | |:|||  |||:|| | | :|:
Db        363 VSGITGVISTILQYSKQAMFEHVANKIHNKIVEWEKNNHGKNYFENGYDARYLANLQDNM 422

Qy        432 KFLSELNKELEAERVIAITQQRWDNNIGELAGITKLGERIKSGKAYADAFEDGKKVEAGS 491
              |||  |||||:||||||||||:||||||:|||:|:||:  ||||| |||||:|||: ::|
Db        423 KFLLNLNKELQAERVIAITQQQWDNNIGDLAGISRLGEKVLSGKAYVDAFEEGKHIKADK 482

Qy        492 NITLDAKTGIIDISNSNGKKTQALHFTSPLLTAGTESRERLTNGKYSYINKLKFGRVKNW 551
                : ||:  ||||:|||    |||  : | :||||  ||| |||:   ||| || ||      || :|
Db        483 LVQLDSANGIIDVSNSGKAKTQHILFRTPLLTPGTEHRERVQTGKYEYITKLNINRVDSW 542

Qy        552 QVTDGEASSKLDFSKVIQRV-------AETEGTDEIGLIVNAKAGNDDIFVGQGKMNIDG 604
              ::||| |||   |  : |:||:                 | |   :|       |:|::||| |      |||
Db        543 KITDGAASSTFDLTNVVQRIGIELDNAGNVTKTKETKIIAKLGEGDDNVFVGSGTTEIDG 602

Qy        605 GDGHDRVFYSKDGGFGNITVDGTSATEAGSYTVNRKVARGDIYHEVVKRQETKVGKRTET 664
              |:|:||| ||:  |  :|   :|:|  |    || ||||||| |     |    |||          || |  |
Db        603 GEGYDRVHYSR-GNYGALTIDATKETEQGSYTVNRFVETGKALHEVTSTHTALVGNREEK 661

Qy        665 IQYRDYELRKVGYGYQSTDNLKSVEEVIGSQFNDVFKGSKFNDIFHSGEGDDLLDGGAGD 724
              |:||  :      :      || :  | ||:|||:||:   ||:||||||||  |:  |:| |  :|| |:
Db        662 IEYR-HSNNQHHAGYYTKDTLKAVEEIIGTSHNDIFKGSKFNDAFNGGDGVDTIDGNDGN 720

Qy        725 DRLFGGKGNDRLSGDEGDDLLDGGSGDDVLNGGAGNDVYIFRKGDGNDTLYDGTGNDKLA 784
              |||||||:|  |  |    |||  :|||  |:|:|:|:|| |:|::::  ||||||  :  |   |||||:
Db        721 DRLFGGKGDDILDGGNGDDFIDGGKGNDLLHGGKGDDIFVHRKGDGNDIITDSDGNDKLS 780

Qy        785 FADANISDIMIERTKEGIIVKRNDHSGSINIPRWY----ITSNLQNYQSNKTDHKIEQLI 840
              |:|:|: |:  |: |   :::  |           :  |  |:             : ||::  |  | |||::|
Db        781 FSDSNLKDLTFEKVKHNLVI-TNSKKEKVTIQNWFREADFAKEVPNYKATK-DEKIEEII 838

Qy        841 GKDGSYITSDQIDKILQDKKDGTVITSQELKKLADENKSQKLSASDIASSLNKLVGSMAL 900
              |::|   |||  |:|  ::      |          ||   || |:  |    :   | |     :: :||:||: |::
Db        839 GQNGERITSKQVDDLI--AKGNGKITQDELSKVVDNYELLKHS-KNVTNSLDKLISSVSA 895

Qy        901 FGTANSVSSNALQPITQPTQGI 922
              | ::|    :   :  |  :    | :
Db        896 FTSSNDSRNVLVAPTSMLDQSL 917
```

```
RESULT 13
5476657-3
;Patent No. 5476657
;    APPLICANT: POTTER, ANDREW A.
;    TITLE OF INVENTION: PASTEURELLA HAEMOLYTICA LEUKOTOXIN
;COMPOSITIONS AND USES THEREOF
;    NUMBER OF SEQUENCES: 8
;    CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/15,537
;       FILING DATE: 09-FEB-1993
;    PRIOR APPLICATION DATA:
;       APPLICATION NUMBER: 504,850
;       FILING DATE: 05-APR-1990
;       APPLICATION NUMBER: 335,018
;       FILING DATE: 07-APR-1989
;SEQ ID NO:3:
;       LENGTH: 926
5476657-3

  Query Match             50.1%;  Score 2326.5;  DB 6;  Length 926;
  Best Local Similarity   50.0%;  Pred. No. 1.5e-173;
  Matches  461;  Conservative  174;  Mismatches  262;  Indels   25;  Gaps   12;

Qy         15 LNSTKSGLKNLYLAIPKD--YDPQKGGTLNDFIKAADELGIARLAEEPNHTETAKKSVDT 72
              |:  |:| | : | ||:: || ::| | | :|||:|||    || |:  ||: |: |
Db          7 LSFPKTGAKKIILYIPQNYQYDTEQGNGLQDLVKAAEELGIEVQREERNNIATAQTSLGT 66

Qy         73 VNQFLSLTQTGIAISATKLEKFLQKHSTNKLAKGLDSVENIDRKLGKASNVLSTLSSFLG 132
               :    : ||: || :|| :::| |||      |   : | | |:| :    ||  ||| : | ||
Db         67 IQTAIGLTERGIVLSAPQIDKLLQK---TKAGQALGSAESIVQNANKAKTVLSGIQSILG 123

Qy        133 TALAGIELDSLIKKGDAAPDALAKASIDLINEIIGNLSQSTQTIEAFSSQLAKLGSTISQ 192
              : |||::||   : : ::      ||||| ::| |  :| |:: | :|:: |  |::: || :
Db        124 SVLAGMDLDEAL-QNNSNQHALAKAGLELTNSLIENIANSVKTLDEFGEQISQFGSKLQN 182

Qy        193 AKGFSNIGNKLQNL-NFSKTNLGLEIITGLLSGISAGFALADKNASTGKKVAAGFELSNQ 251
                ||    :|:||:|:     |  |||::|:||||| :|   |||||||| ||| |||||:||
Db        183 IKGLGTLGDKLKNIGGLDKAGLGLDVISGLLSGATAALVLADKNASTAKKVGAGFELANQ 242

Qy        252 VIGNVTKAISSYVLAQRVAAGLSTTGAVAALITSSIMLAISPLAFMNAADKFNHANALDE 311
              |:||:|:|||:|||:|||||||||||:|| ||||| |||   :  |||||||| |||||||| :|:
Db        243 VVGNITKAVSSYILAQRVAAGLSSTGPVAALIASTVSLAISPLAFAGIADKFNHAKSLES 302

Qy        312 FAKQFRKFGYDGDHLLAEYQRGVGTIEASLTTISTALGAVSAGVSAAAVGSAVGTPIALL 371
              :|::|:| |||||:|||||||| |||:||:| |:||| |:: ||||| ||  : :|||||
Db        303 YAERFKKLGYDGDNLLAEYQRGTGTIDASVTAINTALAAIAGGVSAAAAGSVIASPIALL 362

Qy        372 VAGVTGLISGILEASKQAMFESVANRLQGKILEWEKQNGGQNYFDKGYDSRYAAYLANNL 431
              |:|:||:|| ||: ||||||| |||::  ||:|||  | |:|||: |||:|| | | :|:
Db        363 VSGITGVISTILQYSKQAMFEHVANKIHNKIVEWEKNNHGKNYFENGYDARYLANLQDNM 422

Qy        432 KFLSELNKELEAERVIAITQQRWDNNIGELAGITKLGERIKSGKAYADAFEDGKKVEAGS 491
              |||  |||||:|||||||||||:||||||:|||||::|||:: ||||| ||||:|| ::|
Db        423 KFLLNLNKELQAERVIAITQQQWDNNIGDLAGISRLGEKVLSGKAYVDAFEEGKHIKADK 482

Qy        492 NITLDAKTGIIDISNSNGKKTQALHFTSPLLTAGTESRERLTNGKYSYINKLKFGRVKNW 551
                : ||:  ||||:|||  |||  :  :||||  ||| |||:  ||| ||  ||   || :|
Db        483 LVQLDSANGIIDVSNSGKAKTQHILFRTPLLTPGTEHRERVQTGKYEYITKLNINRVDSW 542

Qy        552 QVTDGEASSKLDFSKVIQRV-------AETEGTDEIGLIVNAKAGNDDIFVGQGKMNIDG 604
              ::||| |||   | :  |:||:              | |   :|      |:|::|||   |||
Db        543 KITDGAASSTFDLTNVVQRIGIELDNAGNVTKTKETKIIAKLGEGDDNVFVGSGTTEIDG 602

Qy        605 GDGHDRVFYSKDGGFGNITVDGTSATEAGSYTVNRKVARGDIYHEVVKRQETKVGKRTET 664
              |:|:||| ||: | :| :|:| |  || |||||| |   |    |||         || | |
Db        603 GEGYDRVHYSR-GNYGALTIDATKETEQGSYTVNRFVETGKALHEVTSTHTALVGNREEK 661

Qy        665 IQYRDYELRKVGYGYQSTDNLKSVEEVIGSQFNDVFKGSKFNDIFHSGEGDDLLDGGAGD 724
              |:|| :    :     || : |  ||:|||:||:  ||:||||||||  |: |:| | :|| |:
Db        662 IEYR-HSNNQHHAGYYTKDTLKAVEEIIGTSHNDIFKGSKFNDAFNGGDGVDTIDGNDGN 720

Qy        725 DRLFGGKGNDRLSGDEGDDLLDGGSGDDVLNGGAGNDVYIFRKGDGNDTLYDGTGNDKLA 784
              ||||||||:| |  |  ||| :|||  |:|:|:|| |:|:::  ||||||  :  |  |||||:
Db        721 DRLFGGKGDDILDGGNGDDFIDGGKGNDLLHGGKGDDIFVHRKGDGNDIITDSDGNDKLS 780

Qy        785 FADANISDIMIERTKEGIIVKRNDHSGSINIPRWY----ITSNLQNYQSNKTDHKIEQLI 840
              |:|:|: |:  |:  |: |  :::  |       :  |  |:  `        : ||:: |  | |||::|
Db        781 FSDSNLKDLTFEKVKHNLVI-TNSKKEKVTIQNWFREADFAKEVPNYKATK-DEKIEEII 838

Qy        841 GKDGSYITSDQIDKILQDKKDGTVITSQELKKLADENKSQKLSASDIASSLNKLVGSMAL 900
              |::|  ||| |:| ::    |       ||  || |:  |   :   |  |   :: :||:||: |::
Db        839 GQNGERITSKQVDDLI--AKGNGKITQDELSKVVDNYELLKHS-KNVTNSLDKLISSVSA 895

Qy        901 FGTANSVSSNALQPITQPTQGI 922
              | ::|    :  :  |  :   | :
Db        896 FTSSNDSRNVLVAPTSMLDQSL 917

RESULT 14
US-08-455-970A-12
; Sequence 12, Application US/08455970A
; Patent No. 5708155
;  GENERAL INFORMATION:
;    APPLICANT:  POTTER, ANDREW A.
;    APPLICANT:  REDMOND, MARK J.
;    APPLICANT:  HUGHES, HUW P.A.
;    TITLE OF INVENTION:  ENHANCED IMMUNOGENICITY USING LEUKOTOXIN
;    TITLE OF INVENTION:  CHIMERAS
;    NUMBER OF SEQUENCES:  15
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  REED & ROBINS
;      STREET:  285 HAMILTON AVENUE, SUITE 200
;      CITY:  PALO ALTO
;      STATE:  CALIFORNIA
;      COUNTRY:  UNITED STATES OF AMERICA
;      ZIP:  94301
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/455,970A
;      FILING DATE:  31-MAY-1995
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/960,932
;      FILING DATE:  14-OCT-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  ROBINS, ROBERTA L.
;      REGISTRATION NUMBER:  33,208
;      REFERENCE/DOCKET NUMBER:  9001-0016.10
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 327-3400
;      TELEFAX:  (415) 327-3231
;  INFORMATION FOR SEQ ID NO:  12:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  936 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-455-970A-12

  Query Match             50.1%;  Score 2326.5;  DB 1;  Length 936;
  Best Local Similarity   50.0%;  Pred. No. 1.6e-173;
  Matches  461;  Conservative  174;  Mismatches  262;  Indels   25;  Gaps   12;

Qy         15 LNSTKSGLKNLYLAIPKD--YDPQKGGTLNDFIKAADELGIARLAEEPNHTETAKKSVDT 72
              |:  |:| | : | ||:: || ::| | | :|||:|||    || |:  ||: |: |
Db          7 LSFPKTGAKKIILYIPQNYQYDTEQGNGLQDLVKAAEELGIEVQREERNNIATAQTSLGT 66

Qy         73 VNQFLSLTQTGIAISATKLEKFLQKHSTNKLAKGLDSVENIDRKLGKASNVLSTLSSFLG 132
               :    : ||: || :|| :::| |||      |   : | | |:| :    ||  ||| : | ||
Db         67 IQTAIGLTERGIVLSAPQIDKLLQK---TKAGQALGSAESIVQNANKAKTVLSGIQSILG 123
```

```
Qy        133 TALAGIELDSLIKKGDAAPDALAKASIDLINEIIGNLSQSTQTIEAFSSQLAKLGSTISQ 192
              : |||::||   : : ::    ||||| ::| | :| |:: | :|:: |  |::: || :
Db        124 SVLAGMDLDEAL-QNNSNQHALAKAGLELTNSLIENIANSVKTLDEFGEQISQFGSKLQN 182

Qy        193 AKGFSNIGNKLQNL-NFSKTNLGLEIITGLLSGISAGFALADKNASTGKKVAAGFELSNQ 251
               ||    :|:||:|:      |  |||::|:||||| :|    ||||||| ||| |||||:||
Db        183 IKGLGTLGDKLKNIGGLDKAGLGLDVISGLLSGATAALVLADKNASTAKKVGAGFELANQ 242

Qy        252 VIGNVTKAISSYVLAQRVAAGLSTTGAVAALITSSIMLAISPLAFMNAADKFNHANALDE 311
              |:||:|||:|||:||||||||||:|| ||||| |::  ||||||||     |||||| :|:
Db        243 VVGNITKAVSSYILAQRVAAGLSSTGPVAALIASTVSLAISPLAFAGIADKFNHAKSLES 302

Qy        312 FAKQFRKFGYDGDHLLAEYQRGVGTIEASLTTISTALGAVSAGVSAAAVGSAVGTPIALL 371
              :|::|:| |||||:||||||||| |||:||:| |:||| |:: ||||| || : :|||||
Db        303 YAERFKKLGYDGDNLLAEYQRGTGTIDASVTAINTALAAIAGGVSAAAAGSVIASPIALL 362

Qy        372 VAGVTGLISGILEASKQAMFESVANRLQGKILEWEKQNGGQNYFDKGYDSRYAAYLANNL 431
              |:|:||:|| ||: |||||| |||::  ||:|||| | |:|||: |||:|| | | :|:
Db        363 VSGITGVISTILQYSKQAMFEHVANKIHNKIVEWEKNNHGKNYFENGYDARYLANLQDNM 422

Qy        432 KFLSELNKELEAERVIAITQQRWDNNIGELAGITKLGERIKSGKAYADAFEDGKKVEAGS 491
              |||  |||||:|||||||||||:||||||:||||::|||:: ||||| ||||:|| ::|
Db        423 KFLLNLNKELQAERVIAITQQQWDNNIGDLAGISRLGEKVLSGKAYVDAFEEGKHIKADK 482

Qy        492 NITLDAKTGIIDISNSNGKKTQALHFTSPLLTAGTESRERLTNGKYSYINKLKFGRVKNW 551
               : ||:  ||||:|||   |||  : | :|||| ||| |||:  ||| ||  ||   || :|
Db        483 LVQLDSANGIIDVSNSGKAKTQHILFRTPLLTPGTEHRERVQTGKYEYITKLNINRVDSW 542

Qy        552 QVTDGEASSKLDFSKVIQRV-------AETEGTDEIGLIVNAKAGNDDIFVGQGKMNIDG 604
              ::||| |||  | : |:||:            | |  :|       |:|::||| |    |||
Db        543 KITDGAASSTFDLTNVVQRIGIELDNAGNVTKTKETKIIAKLGEGDDNVFVGSGTTEIDG 602

Qy        605 GDGHDRVFYSKDGGFGNITVDGTSATEAGSYTVNRKVARGDIYHEVVKRQETKVGKRTET 664
              |:|:||| ||: | :| :|:| |  || ||||||| |  |    |||          || | |
Db        603 GEGYDRVHYSR-GNYGALTIDATKETEQGSYTVNRFVETGKALHEVTSTHTALVGNREEK 661

Qy        665 IQYRDYELRKVGYGYQSTDNLKSVEEVIGSQFNDVFKGSKFNDIFHSGEGDDLLDGGAGD 724
              |:|| :   :    || : | ||:|||:||:  ||:||||||||| |: |:| | :|| |:
Db        662 IEYR-HSNNQHHAGYYTKDTLKAVEEIIGTSHNDIFKGSKFNDAFNGGDGVDTIDGNDGN 720

Qy        725 DRLFGGKGNDRLSGDEGDDLLDGGSGDDVLNGGAGNDVYIFRKGDGNDTLYDGTGNDKLA 784
              ||||||||:| | |  |||:|||  ||| |:|:|:| |::|:| ||||||  : |  |||||:
Db        721 DRLFGGKGDDILDGGNGDDFIDGGKGNDLLHGGKGDDIFVHRKGDGNDIITDSDGNDKLS 780

Qy        785 FADANISDIMIERTKEGIIVKRNDHSGSINIPRWY----ITSNLQNYQSNKTDHKIEQLI 840
              |:|:|: |: |:| |  :::  |      :  |  |:            : ||:: | | |||::|
Db        781 FSDSNLKDLTFEKVKHNLVI-TNSKKEKVTIQNWFREADFAKEVPNYKATK-DEKIEEII 838

Qy        841 GKDGSYITSDQIDKILQDKKDGTVITSQELKKLADENKSQKLSASDIASSLNKLVGSMAL 900
              |::|  ||| |:| ::    |      ||  || |: |   :  | |    :: :||:||: |::
Db        839 GQNGERITSKQVDDLI--AKGNGKITQDELSKVVDNYELLKHS-KNVTNSLDKLISSVSA 895

Qy        901 FGTANSVSSNALQPITQPTQGI 922
              | ::|    :  : | :    | :
Db        896 FTSSNDSRNVLVAPTSMLDQSL 917
```

RESULT 15
US-08-455-970A-10
; Sequence 10, Application US/08455970A
; Patent No. 5708155
; GENERAL INFORMATION:
; APPLICANT: POTTER, ANDREW A.
; APPLICANT: REDMOND, MARK J.
; APPLICANT: HUGHES, HUW P.A.
; TITLE OF INVENTION: ENHANCED IMMUNOGENICITY USING LEUKOTOXIN
; TITLE OF INVENTION: CHIMERAS
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: REED & ROBINS
; STREET: 285 HAMILTON AVENUE, SUITE 200
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/455,970A
; FILING DATE: 31-MAY-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/960,932
; FILING DATE: 14-OCT-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: ROBINS, ROBERTA L.
; REGISTRATION NUMBER: 33,208
; REFERENCE/DOCKET NUMBER: 9001-0016.10
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 327-3400
; TELEFAX: (415) 327-3231
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 943 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-455-970A-10

Query Match 50.1%; Score 2326.5; DB 1; Length 943;
Best Local Similarity 50.0%; Pred. No. 1.6e-173;
Matches 461; Conservative 174; Mismatches 262; Indels 25; Gaps 12;

```
Qy         15 LNSTKSGLKNLYLAIPKD--YDPQKGGTLNDFIKAADELGIARLAEEPNHTETAKKSVDT 72
              |:  |:| | : | ||::  || ::|  | | :||:|||     || |:  ||: |: |
Db          7 LSFPKTGAKKIILYIPQNYQYDTEQGNGLQDLVKAAEELGIEVQREERNNIATAQTSLGT 66

Qy         73 VNQFLSLTQTGIAISATKLEKFLQKHSTNKLAKGLDSVENIDRKLGKASNVLSTLSSFLG 132
               : || :| ||  :::| |||     |  | | |:| :    ||  ||| : ||
Db         67 IQTAIGLTERGIVLSAPQIDKLLQK---TKAGQALGSAESIVQNANKAKTVLSGIQSILG 123

Qy        133 TALAGIELDSLIKKGDAAPDALAKASIDLINEIIGNLSQSTQTIEAFSSQLAKLGSTISQ 192
              : |||::||   : : ::    ||||| ::| | :| |:: | :|:: |  |::: || :
Db        124 SVLAGMDLDEAL-QNNSNQHALAKAGLELTNSLIENIANSVKTLDEFGEQISQFGSKLQN 182

Qy        193 AKGFSNIGNKLQNL-NFSKTNLGLEIITGLLSGISAGFALADKNASTGKKVAAGFELSNQ 251
               ||    :|:||:|:      |  |||::|:||||| :|    ||||||| ||| |||||:||
Db        183 IKGLGTLGDKLKNIGGLDKAGLGLDVISGLLSGATAALVLADKNASTAKKVGAGFELANQ 242

Qy        252 VIGNVTKAISSYVLAQRVAAGLSTTGAVAALITSSIMLAISPLAFMNAADKFNHANALDE 311
              |:||:|||:|||:||||||||||:|| ||||| |::  ||||||||     |||||| :|:
Db        243 VVGNITKAVSSYILAQRVAAGLSSTGPVAALIASTVSLAISPLAFAGIADKFNHAKSLES 302

Qy        312 FAKQFRKFGYDGDHLLAEYQRGVGTIEASLTTISTALGAVSAGVSAAAVGSAVGTPIALL 371
              :|::|:| |||||:||||||||| |||:||:| |:||| |:: ||||| || : :|||||
Db        303 YAERFKKLGYDGDNLLAEYQRGTGTIDASVTAINTALAAIAGGVSAAAAGSVIASPIALL 362

Qy        372 VAGVTGLISGILEASKQAMFESVANRLQGKILEWEKQNGGQNYFDKGYDSRYAAYLANNL 431
              |:|:||:|| ||: |||||| |||::  ||:|||| | |:|||: |||:|| | | :|:
Db        363 VSGITGVISTILQYSKQAMFEHVANKIHNKIVEWEKNNHGKNYFENGYDARYLANLQDNM 422

Qy        432 KFLSELNKELEAERVIAITQQRWDNNIGELAGITKLGERIKSGKAYADAFEDGKKVEAGS 491
              |||  |||||:|||||||||||:||||||:||||::|||:: ||||| ||||:|| ::|
Db        423 KFLLNLNKELQAERVIAITQQQWDNNIGDLAGISRLGEKVLSGKAYVDAFEEGKHIKADK 482

Qy        492 NITLDAKTGIIDISNSNGKKTQALHFTSPLLTAGTESRERLTNGKYSYINKLKFGRVKNW 551
               : ||:  ||||:|||   |||  : | :|||| ||| |||:  ||| ||  ||   || :|
Db        483 LVQLDSANGIIDVSNSGKAKTQHILFRTPLLTPGTEHRERVQTGKYEYITKLNINRVDSW 542

Qy        552 QVTDGEASSKLDFSKVIQRV-------AETEGTDEIGLIVNAKAGNDDIFVGQGKMNIDG 604
              ::||| |||  | : |:||:            | |  :|       |:|::||| |    |||
Db        543 KITDGAASSTFDLTNVVQRIGIELDNAGNVTKTKETKIIAKLGEGDDNVFVGSGTTEIDG 602
```

```
Qy        605 GDGHDRVFYSKDGGFGNITVDGTSATEAGSYTVNRKVARGDIYHEVVKRQETKVGKRTET 664
              |:|:||| ||: | :| :|:| |  || |||||| |  |   |||       || | |
Db        603 GEGYDRVHYSR-GNYGALTIDATKETEQGSYTVNRFVETGKALHEVTSTHTALVGNREEK 661

Qy        665 IQYRDYELRKVGYGYQSTDNLKSVEEVIGSQFNDVFKGSKFNDIFHSGEGDDLLDGGAGD 724
              |:|| :    :     || : | ||:|||:||:  ||:|||||||| |: |:| | :|| |:
Db        662 IEYR-HSNNQHHAGYYTKDTLKAVEEIIGTSHNDIFKGSKFNDAFNGGDGVDTIDGNDGN 720

Qy        725 DRLFGGKGNDRLSGDEGDDLLDGGSGDDVLNGGAGNDVYIFRKGDGNDTLYDGTGNDKLA 784
              ||||||||:| |  |  ||| :||| |:|:|:|| |:|::: |||||| : |  |||||:
Db        721 DRLFGGKGDDILDGGNGDDFIDGGKGNDLLHGGKGDDIFVHRKGDGNDIITDSDGNDKLS 780

Qy        785 FADANISDIMIERTKEGIIVKRNDHSGSINIPRWY----ITSNLQNYQSNKTDHKIEQLI 840
              |:|:|: |:  |: |  ::: |      :  | |:         : ||:: | | |||::|
Db        781 FSDSNLKDLTFEKVKHNLVI-TNSKKEKVTIQNWFREADFAKEVPNYKATK-DEKIEEII 838

Qy        841 GKDGSYITSDQIDKILQDKKDGTVITSQELKKLADENKSQKLSASDIASSLNKLVGSMAL 900
              |::|  ||| |:| ::   |     ||  || |: |   :  | |   :: :||:||: |::
Db        839 GQNGERITSKQVDDLI--AKGNGKITQDELSKVVDNYELLKHS-KNVTNSLDKLISSVSA 895

Qy        901 FGTANSVSSNALQPITQPTQGI 922
              | ::|   :  : | :   | :
Db        896 FTSSNDSRNVLVAPTSMLDQSL 917
```

RESULT 16
US-08-455-970A-14
; Sequence 14, Application US/08455970A
; Patent No. 5708155
; GENERAL INFORMATION:
; APPLICANT: POTTER, ANDREW A.
; APPLICANT: REDMOND, MARK J.
; APPLICANT: HUGHES, HUW P.A.
; TITLE OF INVENTION: ENHANCED IMMUNOGENICITY USING LEUKOTOXIN
; TITLE OF INVENTION: CHIMERAS
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: REED & ROBINS
; STREET: 285 HAMILTON AVENUE, SUITE 200
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/455,970A
; FILING DATE: 31-MAY-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/960,932
; FILING DATE: 14-OCT-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: ROBINS, ROBERTA L.
; REGISTRATION NUMBER: 33,208
; REFERENCE/DOCKET NUMBER: 9001-0016.10
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 327-3400
; TELEFAX: (415) 327-3231
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 951 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-455-970A-14

```
  Query Match             50.1%;  Score 2326.5;  DB 1;  Length 951;
  Best Local Similarity   50.0%;  Pred. No. 1.6e-173;
  Matches  461;  Conservative  174;  Mismatches  262;  Indels   25;  Gaps   12;
```

```
Qy         15 LNSTKSGLKNLYLAIPKD--YDPQKGGTLNDFIKAADELGIARLAEEPNHTETAKKSVDT 72
              |:  |:| | : | ||::  || ::|  | | :|||:|||    || |:  ||: |: |
Db          7 LSFPKTGAKKIILYIPQNYQYDTEQGNGLQDLVKAAEELGIEVQREERNNIATAQTSLGT 66

Qy         73 VNQFLSLTQTGIAISATKLEKFLQKHSTNKLAKGLDSVENIDRKLGKASNVLSTLSSFLG 132
               :  : ||: || :|| :::| |||    |   :  | | |:| :   ||  |||  : | ||
Db         67 IQTAIGLTERGIVLSAPQIDKLLQK---TKAGQALGSAESIVQNANKAKTVLSGIQSILG 123

Qy        133 TALAGIELDSLIKKGDAAPDALAKASIDLINEIIGNLSQSTQTIEAFSSQLAKLGSTISQ 192
              : |||::||  : : ::  ||||  ::| | :| |:: | :|:: |  |::: || :
Db        124 SVLAGMDLDEAL-QNNSNQHALAKAGLELTNSLIENIANSVKTLDEFGEQISQFGSKLQN 182

Qy        193 AKGFSNIGNKLQNL-NFSKTNLGLEIITGLLSGISAGFALADKNASTGKKVAAGFELSNQ 251
               ||  :|:||:|:    |  |||::|:||||| :|  ||||||| ||| |||||:||
Db        183 IKGLGTLGDKLKNIGGLDKAGLGLDVISGLLSGATAALVLADKNASTAKKVGAGFELANQ 242

Qy        252 VIGNVTKAISSYVLAQRVAAGLSTTGAVAALITSSIMLAISPLAFMNAADKFNHANALDE 311
              |:||:|||:|||:||||||||||:|| |||| |::  |||||||||  |||||| :|:
Db        243 VVGNITKAVSSYILAQRVAAGLSSTGPVAALIASTVSLAISPLAFAGIADKFNHAKSLES 302

Qy        312 FAKQFRKFGYDGDHLLAEYQRGVGTIEASLTTISTALGAVSAGVSAAAVGSAVGTPIALL 371
               :|::|:| ||||:||||||| |||:|:|:| |:||| |:: ||||| || :  :|||||
Db        303 YAERFKKLGYDGDNLLAEYQRGTGTIDASVTAINTALAAIAGGVSAAAAGSVIASPIALL 362

Qy        372 VAGVTGLISGILEASKQAMFESVANRLQKILEWEKQNGGQNYFDKGYDSRYAAYLANNL 431
              |:|:||:|| ||: |||||| |||:: ||:||||  | :|||: |||:|| | | :|:
Db        363 VSGITGVISTILQYSKQAMFEHVANKIHNKIVEWEKNNHGKNYFENGYDARYLANLQDNM 422

Qy        432 KFLSELNKELEAERVIAITQQRWDNNIGELAGITKLGERIKSGKAYADAFEDGKKVEAGS 491
              |||  |||||:||||||||||:||||||:||||::|||:: ||||| ||||:|| ::|
Db        423 KFLLNLNKELQAERVIAITQQQWDNNIGDLAGISRLGEKVLSGKAYVDAFEEGKHIKADK 482

Qy        492 NITLDAKTGIIDISNSNGKKTQALHFTSPLLTAGTESRERLTNGKYSYINKLKFGRVKNW 551
               :||:   ||||:|||  |||  :  | :|||| ||| |||:  |||  ||  ||  | :|
Db        483 LVQLDSANGIIDVSNSGKAKTQHILFRTPLLTPGTEHRERVQTGKYEYITKLNINRVDSW 542

Qy        552 QVTDGEASSKLDFSKVIQRV-------AETEGTDEIGLIVNAKAGNDDIFVGQGKMNIDG 604
              ::||| |||  | : |:||:             · | |  :|     |:|::|||  |  |||
Db        543 KITDGAASSTFDLTNVVQRIGIELDNAGNVTKTKETKIIAKLGEGDDNVFVGSGTTEIDG 602

Qy        605 GDGHDRVFYSKDGGFGNITVDGTSATEAGSYTVNRKVARGDIYHEVVKRQETKVGKRTET 664
              |:|:||| ||: | :| :|:| |  || |||||| |  |   |||       || | |
Db        603 GEGYDRVHYSR-GNYGALTIDATKETEQGSYTVNRFVETGKALHEVTSTHTALVGNREEK 661

Qy        665 IQYRDYELRKVGYGYQSTDNLKSVEEVIGSQFNDVFKGSKFNDIFHSGEGDDLLDGGAGD 724
              |:|| :    :     || : | ||:|||:||:  ||:|||||||| |: |:| | :|| |:
Db        662 IEYR-HSNNQHHAGYYTKDTLKAVEEIIGTSHNDIFKGSKFNDAFNGGDGVDTIDGNDGN 720

Qy        725 DRLFGGKGNDRLSGDEGDDLLDGGSGDDVLNGGAGNDVYIFRKGDGNDTLYDGTGNDKLA 784
              ||||||||:| |  |  ||| :||| |:|:|:|| |:|::: |||||| : |  |||||:
Db        721 DRLFGGKGDDILDGGNGDDFIDGGKGNDLLHGGKGDDIFVHRKGDGNDIITDSDGNDKLS 780

Qy        785 FADANISDIMIERTKEGIIVKRNDHSGSINIPRWY----ITSNLQNYQSNKTDHKIEQLI 840
              |:|:|: |:  |: |  ::: |      :  | |:         : ||:: | | |||::|
Db        781 FSDSNLKDLTFEKVKHNLVI-TNSKKEKVTIQNWFREADFAKEVPNYKATK-DEKIEEII 838

Qy        841 GKDGSYITSDQIDKILQDKKDGTVITSQELKKLADENKSQKLSASDIASSLNKLVGSMAL 900
              |::|  ||| |:| ::   |     ||  || |: |   :  | |   :: :||:||: |::
Db        839 GQNGERITSKQVDDLI--AKGNGKITQDELSKVVDNYELLKHS-KNVTNSLDKLISSVSA 895

Qy        901 FGTANSVSSNALQPITQPTQGI 922
              | ::|   :  : | :   | :
Db        896 FTSSNDSRNVLVAPTSMLDQSL 917
```

RESULT 17
US-08-387-156-8
; Sequence 8, Application US/08387156
; Patent No. 5723129
; GENERAL INFORMATION:
; APPLICANT: POTTER, ANDREW A.
; APPLICANT: REDMOND, MARK J.

```
;    APPLICANT:  HUGHES, HUW P.A.
;    TITLE OF INVENTION:  GnRH-LEUKOTOXIN CHIMERAS
;    NUMBER OF SEQUENCES:  28
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  REED & ROBINS
;      STREET:  635 BRYANT STREET
;      CITY:  PALO ALTO
;      STATE:  CALIFORNIA
;      COUNTRY:  UNITED STATES OF AMERICA
;      ZIP:  94301
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/387,156
;      FILING DATE:  10-FEB-1995
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/960,932
;      FILING DATE:  14-OCT-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/779,171
;      FILING DATE:  16-OCT-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  ROBINS, ROBERTA L.
;      REGISTRATION NUMBER:  33,208
;      REFERENCE/DOCKET NUMBER:  9001-0016.21
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 617-8999
;      TELEFAX:  (415) 327-3231
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  977 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-387-156-8

  Query Match            50.1%;  Score 2326.5;  DB 1;  Length 977;
  Best Local Similarity  50.0%;  Pred. No. 1.7e-173;
  Matches  461;  Conservative  174;  Mismatches  262;  Indels   25;  Gaps   12;

Qy         15 LNSTKSGLKNLYLAIPKD--YDPQKGGTLNDFIKAADELGIARLAEEPNHTETAKKSVDT 72
              |:   |:|  | : | ||::  || ::|   |  | :|||:|||    || |:   ||: |: |
Db          7 LSFPKTGAKKIILYIPQNYQYDTEQGNGLQDLVKAAEELGIEVQREERNNIATAQTSLGT 66

Qy         73 VNQFLSLTQTGIAISATKLEKFLQKHSTNKLAKGLDSVENIDRKLGKASNVLSTLSSFLG 132
              :    : ||: || :|| :::| |||      |   : | | |:|  :    ||   ||| : | ||
Db         67 IQTAIGLTERGIVLSAPQIDKLLQK---TKAGQALGSAESIVQNANKAKTVLSGIQSILG 123

Qy        133 TALAGIELDSLIKKGDAAPDALAKASIDLINEIIGNLSQSTQTIEAFSSQLAKLGSTISQ 192
              : |||::||   : : ::    ||||| ::|  | :|  |:: | :|:: |  |::: || :
Db        124 SVLAGMDLDEAL-QNNSNQHALAKAGLELTNSLIENIANSVKTLDEFGEQISQFGSKLQN 182

Qy        193 AKGFSNIGNKLQNL-NFSKTNLGLEIITGLLSGISAGFALADKNASTGKKVAAGFELSNQ 251
               ||     :|:|:|:     |    |||::|:||||  :|    ||||||| ||| |||||:||
Db        183 IKGLGTLGDKLKNIGGLDKAGLGLDVISGLLSGATAALVLADKNASTAKKVGAGFELANQ 242

Qy        252 VIGNVTKAISSYVLAQRVAAGLSTTGAVAALITSSIMLAISPLAFMNAADKFNHANALDE 311
              |:||:|||:|||:||||||||||:|| |||||  |:: |||||||    |||||| :|:
Db        243 VVGNITKAVSSYILAQRVAAGLSSTGPVAALIASTVSLAISPLAFAGIADKFNHAKSLES 302

Qy        312 FAKQFRKFGYDGDHLLAEYQRGVGTIEASLTTISTALGAVSAGVSAAAVGSAVGTPIALL 371
              :|::|:|  |||||:||||||||| |||:||:| |:|||  |:: ||||| || : :|||||
Db        303 YAERFKKLGYDGDNLLAEYQRGTGTIDASVTAINTALAAIAGGVSAAAAGSVIASPIALL 362

Qy        372 VAGVTGLISGILEASKQAMFESVANRLQGKILEWEKQNGGQNYFDKGYDSRYAAYLANNL 431
              |:|:||:|| ||: ||||||| |||::   ||:||||  |  |:|||:  |||:||  |  | :|:
Db        363 VSGITGVISTILQYSKQAMFEHVANKIHNKIVEWEKNNHGKNYFENGYDARYLANLQDNM 422

Qy        432 KFLSELNKELEAERVIAITQQRWDNNIGELAGITKLGERIKSGKAYADAFEDGKKVEAGS 491
              |||  |||||:||||||||||:||||||:|||||::|||:: |||||  ||||:|| ::|
Db        423 KFLLNLNKELQAERVIAITQQQWDNNIGDLAGISRLGEKVLSGKAYVDAFEEGKHIKADK 482

Qy        492 NITLDAKTGIIDISNSNGKKTQALHFTSPLLTAGTESRERLTNGKYSYINKLKFGRVKNW 551
               : ||:  ||||:|||    |||  :  |  :||||  ||| |||:    |||  || ||     || :|
Db        483 LVQLDSANGIIDVSNSGKAKTQHILFRTPLLTPGTEHRERVQTGKYEYITKLNINRVDSW 542

Qy        552 QVTDGEASSKLDFSKVIQRV-------AETEGTDEIGLIVNAKAGNDDIFVGQGKMNIDG 604
              ::|||  |||    |  :  |:|  |:            |  |    :|         |:|::|||  |    |||
Db        543 KITDGAASSTFDLTNVVQRIGIELDNAGNVTKTKETKIIAKLGEGDDNVFVGSGTTEIDG 602

Qy        605 GDGHDRVFYSKDGGFGNITVDGTSATEAGSYTVNRKVARGDIYHEVVKRQETKVGKRTET 664
              |:|:|||  ||:  |  :|  :|:|  |   ||  ||||||| |   |       |||           || |  |
Db        603 GEGYDRVHYSR-GNYGALTIDATKETEQGSYTVNRFVETGKALHEVTSTHTALVGNREEK 661

Qy        665 IQYRDYELRKVGYGYQSTDNLKSVEEVIGSQFNDVFKGSKFNDIFHSGEGDDLLDGGAGD 724
              |:||  :    :    ||  :  |  ||:|||:||:  ||:||||||||  |: |:|  |  :||  |:
Db        662 IEYR-HSNNQHHAGYYTKDTLKAVEEIIGTSHNDIFKGSKFNDAFNGGDGVDTIDGNDGN 720

Qy        725 DRLFGGKGNDRLSGDEGDDLLDGGSGDDVLNGGAGNDVYIFRKGDGNDTLYDGTGNDKLA 784
              |||||||:|  |  |   |||  :|||  |:|:|:||  |:|:::: |||||| :  |   |||||:
Db        721 DRLFGGKGDDILDGGNGDDFIDGGKGNDLLHGGKGDDIFVHRKGDGNDIITDSDGNDKLS 780

Qy        785 FADANISDIMIERTKEGIIVKRNDHSGSINIPRWY----ITSNLQNYQSNKTDHKIEQLI 840
              |:|:|:  |:   |:  |    :::  |        :  |  |:                : ||::  |  | |||::|
Db        781 FSDSNLKDLTFEKVKHNLVI-TNSKKEKVTIQNWFREADFAKEVPNYKATK-DEKIEEII 838

Qy        841 GKDGSYITSDQIDKILQDKKDGTVITSQELKKLADENKSQKLSASDIASSLNKLVGSMAL 900
              |::|   |||  |:|  ::     |       ||   ||   |:  |   :  |   |   ::  :|| :||: |::
Db        839 GQNGERITSKQVDDLI--AKGNGKITQDELSKVVDNYELLKHS-KNVTNSLDKLISSVSA 895

Qy        901 FGTANSVSSNALQPITQPTQGI 922
              | ::|   :  :  | :   | :
Db        896 FTSSNDSRNVLVAPTSMLDQSL 917

RESULT 18
US-08-694-865-8
; Sequence 8, Application US/08694865
; Patent No. 5837268
;  GENERAL INFORMATION:
;    APPLICANT:  POTTER, ANDREW A.
;    APPLICANT:  MANNS, JOHN G.
;    TITLE OF INVENTION:  GnRH-LEUKOTOXIN CHIMERAS
;    NUMBER OF SEQUENCES:  34
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  REED & ROBINS LLP
;      STREET:  285 HAMILTON AVENUE, SUITE 200
;      CITY:  PALO ALTO
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94301
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/694,865
;      FILING DATE:  09-AUG-1996
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  MCCRACKEN, THOMAS P.
;      REGISTRATION NUMBER:  38,548
;      REFERENCE/DOCKET NUMBER:  9001-0016.22
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415)327-3400
;      TELEFAX:  (415)327-3231
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  977 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
```

```
;    MOLECULE TYPE:  protein
US-08-694-865-8

  Query Match             50.1%;  Score 2326.5;  DB 2;  Length 977;
  Best Local Similarity   50.0%;  Pred. No. 1.7e-173;
  Matches  461;  Conservative  174;  Mismatches  262;  Indels   25;  Gaps   12;

Qy         15 LNSTKSGLKNLYLAIPKD--YDPQKGGTLNDFIKAADELGIARLAEEPNHTETAKKSVDT 72
              |:  |:| | : | ||::  || ::|  | | :|||:||||     || |:  ||: |: |
Db          7 LSFPKTGAKKIILYIPQNYQYDTEQGNGLQDLVKAAEELGIEVQREERNNIATAQTSLGT 66

Qy         73 VNQFLSLTQTGIAISATKLEKFLQKHSTNKLAKGLDSVENIDRKLGKASNVLSTLSSFLG 132
              :   : ||: || :|| :::| |||       |  : | | |:| :    ||  ||| : ||
Db         67 IQTAIGLTERGIVLSAPQIDKLLQK---TKAGQALGSAESIVQNANKAKTVLSGIQSILG 123

Qy        133 TALAGIELDSLIKKGDAAPDALAKASIDLINEIIGNLSQSTQTIEAFSSQLAKLGSTISQ 192
              : |||::||  : : ::     ||||| ::| | :| |:: | :|:: |   |::: || :
Db        124 SVLAGMDLDEAL-QNNSNQHALAKAGLELTNSLIENIANSVKTLDEFGEQISQFGSKLQN 182

Qy        193 AKGFSNIGNKLQNL-NFSKTNLGLEIITGLLSGISAGFALADKNASTGKKVAAGFELSNQ 251
               ||    :|:||:|:      |  |||::|:||||| :|   ||||||| ||| |||||:||
Db        183 IKGLGTLGDKLKNIGGLDKAGLGLDVISGLLSGATAALVLADKNASTAKKVGAGFELANQ 242

Qy        252 VIGNVTKAISSYVLAQRVAAGLSTTGAVAALITSSIMLAISPLAFMNAADKFNHANALDE 311
              |:||:|||:|||:|||||||||||:|| ||||| |:: ||||||||    |||||| :|:
Db        243 VVGNITKAVSSYILAQRVAAGLSSTGPVAALIASTVSLAISPLAFAGIADKFNHAKSLES 302

Qy        312 FAKQFRKFGYDGDHLLAEYQRGVGTIEASLTTISTALGAVSAGVSAAAVGSAVGTPIALL 371
               :|::|:| |||||:|||||||| |||:||:| |:||| |:: ||||| || : :||||
Db        303 YAERFKKLGYDGDNLLAEYQRGTGTIDASVTAINTALAAIAGGVSAAAAGSVIASPIALL 362

Qy        372 VAGVTGLISGILEASKQAMFESVANRLQGKILEWEKQNGGQNYFDKGYDSRYAAYLANNL 431
              |:|:||:|| ||: ||||||| |||::  ||:||||  | :|||: |||:|| | | :|:
Db        363 VSGITGVISTILQYSKQAMFEHVANKIHNKIVEWEKNNHGKNYFENGYDARYLANLQDNM 422

Qy        432 KFLSELNKELEAERVIAITQQRWDNNIGELAGITKLGERIKSGKAYADAFEDGKKVEAGS 491
              |||  |||||:||||||||||:||||||:|||||::|||:: ||||| ||||:|| ::|
Db        423 KFLLNLNKELQAERVIAITQQQWDNNIGDLAGISRLGEKVLSGKAYVDAFEEGKHIKADK 482

Qy        492 NITLDAKTGIIDISNSNGKKTQALHFTSPLLTAGTESRERLTNGKYSYINKLKFGRVKNW 551
               : ||:  ||||:|||    ||| :  | :||||  ||| |||:  ||| || ||     || :|
Db        483 LVQLDSANGIIDVSNSGKAKTQHILFRTPLLTPGTEHRERVQTGKYEYITKLNINRVDSW 542

Qy        552 QVTDGEASSKLDFSKVIQRV-------AETEGTDEIGLIVNAKAGNDDIFVGQGKMNIDG 604
              ::||| |||  | : |:||:                   | |   |  |:|::||| |   |||
Db        543 KITDGAASSTFDLTNVVQRIGIELDNAGNVTKTKETKIIAKLGEGDDNVFVGSGTTEIDG 602

Qy        605 GDGHDRVFYSKDGGFGNITVDGTSATEAGSYTVNRKVARGDIYHEVVKRQETKVGKRTET 664
              |:|:||| ||: | :| :|:| |  || |||||||  |   |  |||         || | |
Db        603 GEGYDRVHYSR-GNYGALTIDATKETEQGSYTVNRFVETGKALHEVTSTHTALVGNREEK 661

Qy        665 IQYRDYELRKVGYGYQSTDNLKSVEEVIGSQFNDVFKGSKFNDIFHSGEGDDLLDGGAGD 724
              |:||  :     :   || : | ||:|||:||:  ||:|||||||| |: |:| |  :|| |:
Db        662 IEYR-HSNNQHAGYYTKDTLKAVEEIIGTSHNDIFKGSKFNDAFNGGDGVDTIDGNDGN 720

Qy        725 DRLFGGKGNDRLSGDEGDDLLDGGSGDDVLNGGAGNDVYIFRKGDGNDTLYDGTGNDKLA 784
              |||||||:| | |  ||| :||| |:|:|:|| |:|::: |||||| : |  ||||:
Db        721 DRLFGGKGDDILDGGNGDDFIDGGKGNDLLHGGKGDDIFVHRKGDGNDIITDSDGNDKLS 780

Qy        785 FADANISDIMIERTKEGIIVKRNDHSGSINIPRWY----ITSNLQNYQSNKTDHKIEQLI 840
              |:|:|: |:  |:  |   :::  |     :  |       : ||:: | |  |||::|
Db        781 FSDSNLKDLTFEKVKHNLVI-TNSKKEKVTIQNWFREADFAKEVPNYKATK-DEKIEEII 838

Qy        841 GKDGSYITSDQIDKILQDKKDGTVITSQELKKLADENKSQKLSASDIASSLNKLVGSMAL 900
              |::|  ||| |:|: ::   |      ||  | |  |:  |     ::  :||:||: |::
Db        839 GQNGERITSKQVDDLI--AKGNGKITQDELSKVVDNYELLKHS-KNVTNSLDKLISSVSA 895

Qy        901 FGTANSVSSNALQPITQPTQGI 922
              | ::|    :  : |  :   | :
Db        896 FTSSNDSRNVLVAPTSMLDQSL 917
```

RESULT 19

```
US-08-878-748-8
; Sequence 8, Application US/08878748
; Patent No. 5969126
;  GENERAL INFORMATION:
;    APPLICANT:  POTTER, ANDREW A.
;    APPLICANT:  REDMOND, MARK J.
;    APPLICANT:  HUGHES, HUW P.A.
;    TITLE OF INVENTION:  GnRH-LEUKOTOXIN CHIMERAS
;    NUMBER OF SEQUENCES:  28
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  REED & ROBINS
;      STREET:  635 BRYANT STREET
;      CITY:  PALO ALTO
;      STATE:  CALIFORNIA
;      COUNTRY:  UNITED STATES OF AMERICA
;      ZIP:  94301
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/878,748
;      FILING DATE:  19-JUN-1997
;      CLASSIFICATION:  536
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/387,156
;      FILING DATE:  10-FEB-1995
;      APPLICATION NUMBER:  US 07/960,932
;      FILING DATE:  14-OCT-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/779,171
;      FILING DATE:  16-OCT-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  ROBINS, ROBERTA L.
;      REGISTRATION NUMBER:  33,208
;      REFERENCE/DOCKET NUMBER:  9001-0016.21
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 617-8999
;      TELEFAX:  (415) 327-3231
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  977 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-878-748-8

  Query Match             50.1%;  Score 2326.5;  DB 2;  Length 977;
  Best Local Similarity   50.0%;  Pred. No. 1.7e-173;
  Matches  461;  Conservative  174;  Mismatches  262;  Indels   25;  Gaps   12;

Qy         15 LNSTKSGLKNLYLAIPKD--YDPQKGGTLNDFIKAADELGIARLAEEPNHTETAKKSVDT 72
              |:  |:| | : | ||::  || ::|  | | :|||:||||     || |:  ||: |: |
Db          7 LSFPKTGAKKIILYIPQNYQYDTEQGNGLQDLVKAAEELGIEVQREERNNIATAQTSLGT 66

Qy         73 VNQFLSLTQTGIAISATKLEKFLQKHSTNKLAKGLDSVENIDRKLGKASNVLSTLSSFLG 132
              :   : ||: || :|| :::| |||       |  : | | |:| :    ||  ||| : ||
Db         67 IQTAIGLTERGIVLSAPQIDKLLQK---TKAGQALGSAESIVQNANKAKTVLSGIQSILG 123

Qy        133 TALAGIELDSLIKKGDAAPDALAKASIDLINEIIGNLSQSTQTIEAFSSQLAKLGSTISQ 192
              : |||::||  : : ::     ||||| ::| | :| |:: | :|:: |   |::: || :
Db        124 SVLAGMDLDEAL-QNNSNQHALAKAGLELTNSLIENIANSVKTLDEFGEQISQFGSKLQN 182

Qy        193 AKGFSNIGNKLQNL-NFSKTNLGLEIITGLLSGISAGFALADKNASTGKKVAAGFELSNQ 251
               ||    :|:||:|:      |  |||::|:||||| :|   ||||||| ||| |||||:||
Db        183 IKGLGTLGDKLKNIGGLDKAGLGLDVISGLLSGATAALVLADKNASTAKKVGAGFELANQ 242

Qy        252 VIGNVTKAISSYVLAQRVAAGLSTTGAVAALITSSIMLAISPLAFMNAADKFNHANALDE 311
              |:||:|||:|||:|||||||||||:|| ||||| |:: ||||||||    |||||| :|:
Db        243 VVGNITKAVSSYILAQRVAAGLSSTGPVAALIASTVSLAISPLAFAGIADKFNHAKSLES 302

Qy        312 FAKQFRKFGYDGDHLLAEYQRGVGTIEASLTTISTALGAVSAGVSAAAVGSAVGTPIALL 371
```

```
Db   303 YAERFKKLGYDGDNLLAEYQRGTGTIDASVTAINTALAAIAGGVSAAAAGSVIASPIALL 362

Qy   372 VAGVTGLISGILEASKQAMFESVANRLQGKILEWEKQNGGQNYFDKGYDSRYAAYLANNL 431
Db   363 VSGITGVISTILQYSKQAMFEHVANKIHNKIVEWEKNNHGKNYFENGYDARYLANLQDNM 422

Qy   432 KFLSELNKELEAERVIAITQQRWDNNIGELAGITKLGERIKSGKAYADAFEDGKKVEAGS 491
Db   423 KFLLNLNKELQAERVIAITQQQWDNNIGDLAGISRLGEKVLSGKAYVDAFEEGKHIKADK 482

Qy   492 NITLDAKTGIIDISNSNGKKTQALHFTSPLLTAGTESRERLTNGKYSYINKLKFGRVKNW 551
Db   483 LVQLDSANGIIDVSNSGKAKTQHILFRTPLLTPGTEHRERVQTGKYEYITKLNINRVDSW 542

Qy   552 QVTDGEASSKLDFSKVIQRV-------AETEGTDEIGLIVNAKAGNDDIFVGQGKMNIDG 604
Db   543 KITDGAASSTFDLTNVVQRIGIELDNAGNVTKTKETKIIAKLGEGDDNVFVGSGTTEIDG 602

Qy   605 GDGHDRVFYSKDGGFGNITVDGTSATEAGSYTVNRKVARGDIYHEVVKRQETKVGKRTET 664
Db   603 GEGYDRVHYSR-GNYGALTIDATKETEQGSYTVNRFVETGKALHEVTSTHTALVGNREEK 661

Qy   665 IQYRDYELRKVGYGYQSTDNLKSVEEVIGSQFNDVFKGSKFNDIFHSGEGDDLLDGGAGD 724
Db   662 IEYR-HSNNQHHAGYYTKDTLKAVEEIIGTSHNDIFKGSKFNDAFNGGDGVDTIDGNDGN 720

Qy   725 DRLFGGKGNDRLSGDEGDDLLDGGSGDDVLNGGAGNDVYIFRKGDGNDTLYDGTGNDKLA 784
Db   721 DRLFGGKGDDILDGGNGDDFIDGGKGNDLLHGGKGDDIFVHRKGDGNDIITDSDGNDKLS 780

Qy   785 FADANISDIMIERTKEGIIVKRNDHSGSINIPRWY----ITSNLQNYQSNKTDHKIEQLI 840
Db   781 FSDSNLKDLTFEKVKHNLVI-TNSKKEKVTIQNWFREADFAKEVPNYKATK-DEKIEEII 838

Qy   841 GKDGSYITSDQIDKILQDKKDGTVITSQELKKLADENKSQKLSASDIASSLNKLVGSMAL 900
Db   839 GQNGERITSKQVDDLI--AKGNGKITQDELSKVVDNYELLKHS-KNVTNSLDKLISSVSA 895

Qy   901 FGTANSVSSNALQPITQPTQGI 922
Db   896 FTSSNDSRNVLVAPTSMLDQSL 917
```

RESULT 20
US-09-124-491-8
; Sequence 8, Application US/09124491
; Patent No. 6022960
; GENERAL INFORMATION:
; APPLICANT: POTTER, ANDREW A.
; APPLICANT: MANNS, JOHN G.
; TITLE OF INVENTION: GnRH-LEUKOTOXIN CHIMERAS
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: REED & ROBINS LLP
; STREET: 285 HAMILTON AVENUE, SUITE 200
; CITY: PALO ALTO
; STATE: CA
; COUNTRY: USA
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/124,491
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/694,865
; FILING DATE: 09-AUG-1996
; APPLICATION NUMBER: US 08/387,156
; FILING DATE: 10-FEB-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/960,932
; FILING DATE: 14-OCT-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/779,171
; FILING DATE: 16-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: MCCRACKEN, THOMAS P.
; REGISTRATION NUMBER: 38,548
; REFERENCE/DOCKET NUMBER: 9001-0016.22
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415)327-3400
; TELEFAX: (415)327-3231
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 977 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-124-491-8

```
  Query Match             50.1%;  Score 2326.5;  DB 3;  Length 977;
  Best Local Similarity   50.0%;  Pred. No. 1.7e-173;
  Matches  461;  Conservative  174;  Mismatches  262;  Indels   25;  Gaps   12;

Qy    15 LNSTKSGLKNLYLAIPKD--YDPQKGGTLNDFIKAADELGIARLAEEPNHTETAKKSVDT 72
Db     7 LSFPKTGAKKIILYIPQNYQYDTEQGNGLQDLVKAAEELGIEVQREERNNIATAQTSLGT 66

Qy    73 VNQFLSLTQTGIAISATKLEKFLQKHSTNKLAKGLDSVENIDRKLGKASNVLSTLSSFLG 132
Db    67 IQTAIGLTERGIVLSAPQIDKLLQK---TKAGQALGSAESIVQNANKAKTVLSGIQSILG 123

Qy   133 TALAGIELDSLIKKGDAAPDALAKASIDLINEIIGNLSQSTQTIEAFSSQLAKLGSTISQ 192
Db   124 SVLAGMDLDEAL-QNNSNQHALAKAGLELTNSLIENIANSVKTLDEFGEQISQFGSKLQN 182

Qy   193 AKGFSNIGNKLQNL-NFSKTNLGLEIITGLLSGISAGFALADKNASTGKKVAAGFELSNQ 251
Db   183 IKGLGTLGDKLKNIGGLDKAGLGLDVISGLLSGATAALVLADKNASTAKKVGAGFELANQ 242

Qy   252 VIGNVTKAISSYVLAQRVAAGLSTTGAVAALITSSIMLAISPLAFMNAADKFNHANALDE 311
Db   243 VVGNITKAVSSYILAQRVAAGLSSTGPVAALIASTVSLAISPLAFAGIADKFNHAKSLES 302

Qy   312 FAKQFRKFGYDGDHLLAEYQRGVGTIEASLTTISTALGAVSAGVSAAAVGSAVGTPIALL 371
Db   303 YAERFKKLGYDGDNLLAEYQRGTGTIDASVTAINTALAAIAGGVSAAAAGSVIASPIALL 362

Qy   372 VAGVTGLISGILEASKQAMFESVANRLQGKILEWEKQNGGQNYFDKGYDSRYAAYLANNL 431
Db   363 VSGITGVISTILQYSKQAMFEHVANKIHNKIVEWEKNNHGKNYFENGYDARYLANLQDNM 422

Qy   432 KFLSELNKELEAERVIAITQQRWDNNIGELAGITKLGERIKSGKAYADAFEDGKKVEAGS 491
Db   423 KFLLNLNKELQAERVIAITQQQWDNNIGDLAGISRLGEKVLSGKAYVDAFEEGKHIKADK 482

Qy   492 NITLDAKTGIIDISNSNGKKTQALHFTSPLLTAGTESRERLTNGKYSYINKLKFGRVKNW 551
Db   483 LVQLDSANGIIDVSNSGKAKTQHILFRTPLLTPGTEHRERVQTGKYEYITKLNINRVDSW 542

Qy   552 QVTDGEASSKLDFSKVIQRV-------AETEGTDEIGLIVNAKAGNDDIFVGQGKMNIDG 604
Db   543 KITDGAASSTFDLTNVVQRIGIELDNAGNVTKTKETKIIAKLGEGDDNVFVGSGTTEIDG 602

Qy   605 GDGHDRVFYSKDGGFGNITVDGTSATEAGSYTVNRKVARGDIYHEVVKRQETKVGKRTET 664
Db   603 GEGYDRVHYSR-GNYGALTIDATKETEQGSYTVNRFVETGKALHEVTSTHTALVGNREEK 661

Qy   665 IQYRDYELRKVGYGYQSTDNLKSVEEVIGSQFNDVFKGSKFNDIFHSGEGDDLLDGGAGD 724
Db   662 IEYR-HSNNQHHAGYYTKDTLKAVEEIIGTSHNDIFKGSKFNDAFNGGDGVDTIDGNDGN 720
```

```
Qy        725 DRLFGGKGNDRLSGDEGDDLLDGGSGDDVLNGGAGNDVYIFRKGDGNDTLYDGTGNDKLA 784
              |||||||:|  | |  ||| :|||  |:|:|:|| |:|:::  ||||||  :  |   |||||:
Db        721 DRLFGGKGDDILDGGNGDDFIDGGKGNDLLHGGKGDDIFVHRKGDGNDIITDSDGNDKLS 780

Qy        785 FADANISDIMIERTKEGIIVKRNDHSGSINIPRWY----ITSNLQNYQSNKTDHKIEQLI 840
              |:|:|: |:   |: |  :::  |        :  |  |:           : ||:: | | |||::|
Db        781 FSDSNLKDLTFEKVKHNLVI-TNSKKEKVTIQNWFREADFAKEVPNYKATK-DEKIEEII 838

Qy        841 GKDGSYITSDQIDKILQDKKDGTVITSQELKKLADENKSQKLSASDIASSLNKLVGSMAL 900
              |::|  ||| |:|  ::     |      ||   || |:  |    :   |  |    :: :||:||: |::
Db        839 GQNGERITSKQVDDLI--AKGNGKITQDELSKVVDNYELLKHS-KNVTNSLDKLISSVSA 895

Qy        901 FGTANSVSSNALQPITQPTQGI 922
              | ::|    :   :  | :    | :
Db        896 FTSSNDSRNVLVAPTSMLDQSL 917
```

RESULT 21
US-09-383-912-8
; Sequence 8, Application US/09383912
; Patent No. 6521746
; GENERAL INFORMATION:
; APPLICANT: POTTER, ANDREW A.
; APPLICANT: MANNS, JOHN G.
; TITLE OF INVENTION: GnRH-LEUKOTOXIN CHIMERAS
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: REED & ROBINS LLP
; STREET: 285 HAMILTON AVENUE, SUITE 200
; CITY: PALO ALTO
; STATE: CA
; COUNTRY: USA
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/383,912
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/694,865
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: MCCRACKEN, THOMAS P.
; REGISTRATION NUMBER: 38,548
; REFERENCE/DOCKET NUMBER: 9001-0016.22
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415)327-3400
; TELEFAX: (415)327-3231
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 977 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-383-912-8

```
  Query Match             50.1%;  Score 2326.5;  DB 4;  Length 977;
  Best Local Similarity   50.0%;  Pred. No. 1.7e-173;
  Matches  461;  Conservative  174;  Mismatches  262;  Indels   25;  Gaps   12;

Qy         15 LNSTKSGLKNLYLAIPKD--YDPQKGGTLNDFIKAADELGIARLAEEPNHTETAKKSVDT 72
              |:  |:| | : | ||::  || ::|  | | :|||:||||        ||  |:   ||: |: |
Db          7 LSFPKTGAKKIILYIPQNYQYDTEQGNGLQDLVKAAEELGIEVQREERNNIATAQTSLGT 66

Qy         73 VNQFLSLTQTGIAISATKLEKFLQKHSTNKLAKGLDSVENIDRKLGKASNVLSTLSSFLG 132
              :   : ||: || :|| :::| |||        |   : | | |: |    ||  |||  : |  ||
Db         67 IQTAIGLTERGIVLSAPQIDKLLQK---TKAGQALGSAESIVQNANKAKTVLSGIQSILG 123

Qy        133 TALAGIELDSLIKKGDAAPDALAKASIDLINEIIGNLSQSTQTIEAFSSQLAKLGSTISQ 192
              : |||::||  : : ::     ||||| ::| | :| |:: | :|:: |  |::: || :
Db        124 SVLAGMDLDEAL-QNNSNQHALAKAGLELTNSLIENIANSVKTLDEFGEQISQFGSKLQN 182

Qy        193 AKGFSNIGNKLQNL-NFSKTNLGLEIITGLLSGISAGFALADKNASTGKKVAAGFELSNQ 251
              ||   :|:||:|:       |  |||::|:||||  :|   |||||||  |||  ||||||:||
Db        183 IKGLGTLGDKLKNIGGLDKAGLGLDVISGLLSGATAALVLADKNASTAKKVGAGFELANQ 242

Qy        252 VIGNVTKAISSYVLAQRVAAGLSTTGAVAALITSSIMLAISPLAFMNAADKFNHANALDE 311
              |:||:|||:|||:|||||||||||:|| |||||  |::  ||||||||      ||||||  :|:
Db        243 VVGNITKAVSSYILAQRVAAGLSSTGPVAALIASTVSLAISPLAFAGIADKFNHAKSLES 302

Qy        312 FAKQFRKFGYDGDHLLAEYQRGVGTIEASLTTISTALGAVSAGVSAAAVGSAVGTPIALL 371
              :|::|:| |||||:||||||||  |||:||:| |::||| |::  ||||| ||  :  :|||||
Db        303 YAERFKKLGYDGDNLLAEYQRGTGTIDASVTAINTALAAIAGGVSAAAAGSVIASPIALL 362

Qy        372 VAGVTGLISGILEASKQAMFESVANRLQGKILEWEKQNGGQNYFDKGYDSRYAAYLANNL 431
              |:|:||:|| ||:  ||||||  |||::   ||:||||  |  |:|||:  |||:|| |  |  :|:
Db        363 VSGITGVISTILQYSKQAMFEHVANKIHNKIVEWEKNNHGKNYFENGYDARYLANLQDNM 422

Qy        432 KFLSELNKELEAERVIAITQQRWDNNIGELAGITKLGERIKSGKAYADAFEDGKKVEAGS 491
              |||  |||||:|||||||||||:|||||:|||::  ||||  |||||:||  ::|
Db        423 KFLLNLNKELQAERVIAITQQQWDNNIGDLAGISRLGEKVLSGKAYVDAFEEGKHIKADK 482

Qy        492 NITLDAKTGIIDISNSNGKKTQALHFTSPLLTAGTESRERLTNGKYSYINKLKFGRVKNW 551
              : ||:  ||||:|||   |||  :  |  :|||| ||| |||:  ||| || ||    || :|
Db        483 LVQLDSANGIIDVSNSGKAKTQHILFRTPLLTPGTEHRERVQTGKYEYITKLNINRVDSW 542

Qy        552 QVTDGEASSKLDFSKVIQRV-------AETEGTDEIGLIVNAKAGNDDIFVGQGKMNIDG 604
              ::||| |||  |  : |:||:              |  |   :|    |:|::|||  |   |||
Db        543 KITDGAASSTFDLTNVVQRIGIELDNAGNVTKTKETKIIAKLGEGDDNVFVGSGTTEIDG 602

Qy        605 GDGHDRVFYSKDGGFGNITVDGTSATEAGSYTVNRKVARGDIYHEVVKRQETKVGKRTET 664
              |:|:||| ||:  | :| :|:| |   || |||||||  |   |   |||         || | |
Db        603 GEGYDRVHYSR-GNYGALTIDATKETEQGSYTVNRFVETGKALHEVTSTHTALVGNREEK 661

Qy        665 IQYRDYELRKVGYGYQSTDNLKSVEEVIGSQFNDVFKGSKFNDIFHSGEGDDLLDGGAGD 724
              |:||  :            || :  ||:|:||||:|||::  ||:||||||||  |: |:| :|| |:
Db        662 IEYR-HSNNQHHAGYYTKDTLKAVEEIIGTSHNDIFKGSKFNDAFNGGDGVDTIDGNDGN 720

Qy        725 DRLFGGKGNDRLSGDEGDDLLDGGSGDDVLNGGAGNDVYIFRKGDGNDTLYDGTGNDKLA 784
              |||||||:|  | |  ||| :|||  |:|:|:|| |:|:::  ||||||  :  |   |||||:
Db        721 DRLFGGKGDDILDGGNGDDFIDGGKGNDLLHGGKGDDIFVHRKGDGNDIITDSDGNDKLS 780

Qy        785 FADANISDIMIERTKEGIIVKRNDHSGSINIPRWY----ITSNLQNYQSNKTDHKIEQLI 840
              |:|:|: |:   |: |  :::  |        :  |  |:           : ||:: | | |||::|
Db        781 FSDSNLKDLTFEKVKHNLVI-TNSKKEKVTIQNWFREADFAKEVPNYKATK-DEKIEEII 838

Qy        841 GKDGSYITSDQIDKILQDKKDGTVITSQELKKLADENKSQKLSASDIASSLNKLVGSMAL 900
              |::|  ||| |:|  ::     |      ||   || |:  |    :   |  |    :: :||:||: |::
Db        839 GQNGERITSKQVDDLI--AKGNGKITQDELSKVVDNYELLKHS-KNVTNSLDKLISSVSA 895

Qy        901 FGTANSVSSNALQPITQPTQGI 922
              | ::|    :   :  | :    | :
Db        896 FTSSNDSRNVLVAPTSMLDQSL 917
```

RESULT 22
US-07-777-715-9
; Sequence 9, Application US/07777715
; Patent No. 5273889
; GENERAL INFORMATION:
; APPLICANT: Potter, Andrew
; APPLICANT: Campos, Manuel
; APPLICANT: Hughes, Huw P.A.
; TITLE OF INVENTION: CYTOKINE-LEUKOTOXIN GENE FUSIONS AND
; TITLE OF INVENTION: USES THEREOF
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morrison & Foerster
; STREET: 545 Middlefield Road, Suite 200
; CITY: Menlo Park
; STATE: California

```
;     COUNTRY:  USA
;     ZIP:  94025
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE:  Floppy disk
;     COMPUTER:  IBM PC compatible
;     OPERATING SYSTEM:  PC-DOS/MS-DOS
;     SOFTWARE:  PatentIn Release #1.0, Version #1.25
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER:  US/07/777,715
;     FILING DATE:  19911016
;     CLASSIFICATION:  424
;   ATTORNEY/AGENT INFORMATION:
;     NAME:  Robins, Roberta L.
;     REGISTRATION NUMBER:  33,208
;     REFERENCE/DOCKET NUMBER:  29310-2001320
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE:  415-327-7250
;     TELEFAX:  415-327-2951
;     TELEX:  706141
; INFORMATION FOR SEQ ID NO:  9:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH:  1069 amino acids
;     TYPE:  AMINO ACID
;     TOPOLOGY:  linear
;   MOLECULE TYPE:  protein
US-07-777-715-9

  Query Match             50.1%;  Score 2326.5;  DB 1;  Length 1069;
  Best Local Similarity   50.0%;  Pred. No. 1.9e-173;
  Matches  461;  Conservative  174;  Mismatches  262;  Indels   25;  Gaps   12;

Qy         15 LNSTKSGLKNLYLAIPKD--YDPQKGGTLNDFIKAADELGIARLAEEPNHTETAKKSVDT 72
              |:  |:| | : | ||::  || ::|  | | :|||:||||       || |:  ||: |: |
Db          7 LSFPKTGAKKIILYIPQNYQYDTEQGNGLQDLVKAAEELGIEVQREERNNIATAQTSLGT 66

Qy         73 VNQFLSLTQTGIAISATKLEKFLQKHSTNKLAKGLDSVENIDRKLGKASNVLSTLSSFLG 132
              :   : ||: || :|| :::| |||      |  : | | |:| :    ||   ||| : | ||
Db         67 IQTAIGLTERGIVLSAPQIDKLLQK---TKAGQALGSAESIVQNANKAKTVLSGIQSILG 123

Qy        133 TALAGIELDSLIKKGDAAPDALAKASIDLINEIIGNLSQSTQTIEAFSSQLAKLGSTISQ 192
              : |||::||   : : ::     ||||| ::| | :| |:: | :|:: |   |::: || :
Db        124 SVLAGMDLDEAL-QNNSNQHALAKAGLELTNSLIENIANSVKTLDEFGEQISQFGSKLQN 182

Qy        193 AKGFSNIGNKLQNL-NFSKTNLGLEIITGLLSGISAGFALADKNASTGKKVAAGFELSNQ 251
              ||    :|:||:|:    |    |||::|:||||| :|     |||||||| ||| |||||:||
Db        183 IKGLGTLGDKLKNIGGLDKAGLGLDVISGLLSGATAALVLADKNASTAKKVGAGFELANQ 242

Qy        252 VIGNVTKAISSYVLAQRVAAGLSTTGAVAALITSSIMLAISPLAFMNAADKFNHANALDE 311
              |:||:|||:|||:|||||||||:|| |||||  |:: ||||||||    ||||||| :|:
Db        243 VVGNITKAVSSYILAQRVAAGLSSTGPVAALIASTVSLAISPLAFAGIADKFNHAKSLES 302

Qy        312 FAKQFRKFGYDGDHLLAEYQRGVGTIEASLTTISTALGAVSAGVSAAAVGSAVGTPIALL 371
              :|::|:| |||||:||||||||| |||:|:|  |:||| |:: ||||||  || : :|||||
Db        303 YAERFKKLGYDGDNLLAEYQRGTGTIDASVTAINTALAAIAGGVSAAAAGSVIASPIALL 362

Qy        372 VAGVTGLISGILEASKQAMFESVANRLQGKILEWEKQNGGQNYFDKGYDSRYAAYLANNL 431
              |:|:||:|| ||: |||||| |||::  ||:||||  | |:|||: |||:|| | |  :|:
Db        363 VSGITGVISTILQYSKQAMFEHVANKIHNKIVEWEKNNHGKNYFENGYDARYLANLQDNM 422

Qy        432 KFLSELNKELEAERVIAITQQRWDNNIGELAGITKLGERIKSGKAYADAFEDGKKVEAGS 491
              |||  |||||:|||||||||:||||||:|||||||::|||:: |||| ||||:|| ::|
Db        423 KFLLNLNKELQAERVIAITQQQWDNNIGDLAGISRLGEKVLSGKAYVDAFEEGKHIKADK 482

Qy        492 NITLDAKTGIIDISNSNGKKTQALHFTSPLLTAGTESRERLTNGKYSYINKLKFGRVKNW 551
               : ||:  ||||:|||   |||  | : :|||| ||| |||:  ||| || ||    || :|
Db        483 LVQLDSANGIIDVSNSGKAKTQHILFRTPLLTPGTEHRERVQTGKYEYITKLNINRVDSW 542

Qy        552 QVTDGEASSKLDFSKVIQRV-------AETEGTDEIGLIVNAKAGNDDIFVGQGKMNIDG 604
              ::||| |||  | : |:||:              | |  :|      |:|::||| |   |||
Db        543 KITDGAASSTFDLTNVVQRIGIELDNAGNVTKTKETKIIAKLGEGDDNVFVGSGTTEIDG 602

Qy        605 GDGHDRVFYSKDGGFGNITVDGTSATEAGSYTVNRKVARGDIYHEVVKRQETKVGKRTET 664
              |:|:||| ||: | :| :|:| |  |  || ||||||| |   |   |||        || | |
Db        603 GEGYDRVHYSR-GNYGALTIDATKETEQGSYTVNRFVETGKALHEVTSTHTALVGNREEK 661

Qy        665 IQYRDYELRKVGYGYQSTDNLKSVEEVIGSQFNDVFKGSKFNDIFHSGEGDDLLDGGAGD 724
              |:|| :    :     || : | ||:|||:||:   ||:||||||||| |: |:| | :|| |:
Db        662 IEYR-HSNNQHHAGYYTKDTLKAVEEIIGTSHNDIFKGSKFNDAFNGGDGVDTIDGNDGN 720

Qy        725 DRLFGGKGNDRLSGDEGDDLLDGGSGDDVLNGGAGNDVYIFRKGDGNDTLYDGTGNDKLA 784
              ||||||||:| | |   ||| :|||  |:|:|:|| |:|::: ||||||  :  |   |||||:
Db        721 DRLFGGKGDDILDGGNGDDFIDGGKGNDLLHGGKGDDIFVHRKGDGNDIITDSDGNDKLS 780

Qy        785 FADANISDIMIERTKEGIIVKRNDHSGSINIPRWY----ITSNLQNYQSNKTDHKIEQLI 840
              |:|:|: |:  |: |  :::  |        : |  |:            : ||::  | | |||::|
Db        781 FSDSNLKDLTFEKVKHNLVI-TNSKKEKVTIQNWFREADFAKEVPNYKATK-DEKIEEII 838

Qy        841 GKDGSYITSDQIDKILQDKKDGTVITSQELKKLADENKSQKLSASDIASSLNKLVGSMAL 900
              |::|   |||  |:|  ::    |       ||    ||  |:  |   :    |  ::  :||:||: |::
Db        839 GQNGERITSKQVDDLI--AKGNGKITQDELSKVVDNYELLKHS-KNVTNSLDKLISSVSA 895

Qy        901 FGTANSVSSNALQPITQPTQGI 922
              | ::|    :  : | :   | :
Db        896 FTSSNDSRNVLVAPTSMLDQSL 917
```

```
RESULT 23
US-08-170-126-4
; Sequence 4, Application US/08170126
; Patent No. 5594107
;  GENERAL INFORMATION:
;    APPLICANT:  POTTER, ANDREW
;    APPLICANT:  CAMPOS, MANUEL
;    APPLICANT:  HUGHES, HUW P.A.
;    TITLE OF INVENTION:  CYTOKINE-CYTOTOXIN GENE FUSIONS AND USES
;    TITLE OF INVENTION:  THEREOF
;    NUMBER OF SEQUENCES:  6
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  REED & ROBINS
;      STREET:  635 BRYANT STREET
;      CITY:  PALO ALTO
;      STATE:  CALIFORNIA
;      COUNTRY:  UNITED STATES OF AMERICA
;      ZIP:  94301
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/170,126
;      FILING DATE:  20-DEC-1993
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/777,715
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/571,301
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  ROBINS, ROBERTA L.
;      REGISTRATION NUMBER:  33,208
;      REFERENCE/DOCKET NUMBER:  9000-0013.21
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 617-8999
;      TELEFAX:  (415) 327-3231
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1069 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-170-126-4

  Query Match             50.1%;  Score 2326.5;  DB 1;  Length 1069;
  Best Local Similarity   50.0%;  Pred. No. 1.9e-173;
  Matches  461;  Conservative  174;  Mismatches  262;  Indels   25;  Gaps   12;
```

```
Qy         15 LNSTKSGLKNLYLAIPKD--YDPQKGGTLNDFIKAADELGIARLAEEPNHTETAKKSVDT 72
              |:  |:| | : | ||::  || ::|  | | :|||:||||     || |:  ||: |: |
Db          7 LSFPKTGAKKIILYIPQNYQYDTEQGNGLQDLVKAAEELGIEVQREERNNIATAQTSLGT 66

Qy         73 VNQFLSLTQTGIAISATKLEKFLQKHSTNKLAKGLDSVENIDRKLGKASNVLSTLSSFLG 132
              :   : ||: || :|| :::| |||     |  : | | |:| :    ||  ||| : | ||
Db         67 IQTAIGLTERGIVLSAPQIDKLLQK---TKAGQALGSAESIVQNANKAKTVLSGIQSILG 123

Qy        133 TALAGIELDSLIKKGDAAPDALAKASIDLINEIIGNLSQSTQTIEAFSSQLAKLGSTISQ 192
              : |||::||   : : ::   ||||| ::| | :| |:: | :|:: |  |::: || :
Db        124 SVLAGMDLDEAL-QNNSNQHALAKAGLELTNSLIENIANSVKTLDEFGEQISQFGSKLQN 182

Qy        193 AKGFSNIGNKLQNL-NFSKTNLGLEIITGLLSGISAGFALADKNASTGKKVAAGFELSNQ 251
               ||    :|:||:|:    |  |||::|:||||| :|  ||||||| ||| |||||:||
Db        183 IKGLGTLGDKLKNIGGLDKAGLGLDVISGLLSGATAALVLADKNASTAKKVGAGFELANQ 242

Qy        252 VIGNVTKAISSYVLAQRVAAGLSTTGAVAALITSSIMLAISPLAFMNAADKFNHANALDE 311
              |:||:|||:|||:||||||||||:|| ||||| |:: ||||||||     |||||| :|:
Db        243 VVGNITKAVSSYILAQRVAAGLSSTGPVAALIASTVSLAISPLAFAGIADKFNHAKSLES 302

Qy        312 FAKQFRKFGYDGDHLLAEYQRGVGTIEASLTTISTALGAVSAGVSAAAVGSAVGTPIALL 371
              :|::|:| |||||:|||||||| |||:||:| |:||| |:: ||||| ||  : :|||||
Db        303 YAERFKKLGYDGDNLLAEYQRGTGTIDASVTAINTALAAIAGGVSAAAAGSVIASPIALL 362

Qy        372 VAGVTGLISGILEASKQAMFESVANRLQGKILEWEKQNGGQNYFDKGYDSRYAAYLANNL 431
              |:|:||:|| ||: |||||| |||::  ||:|||| | |:|||: |||:|| | | :|:
Db        363 VSGITGVISTILQYSKQAMFEHVANKIHNKIVEWEKNNHGKNYFENGYDARYLANLQDNM 422

Qy        432 KFLSELNKELEAERVIAITQQRWDNNIGELAGITKLGERIKSGKAYADAFEDGKKVEAGS 491
              |||  |||||:||||||||||:||||||:||||::|||:: |||||  ||||:|| ::|
Db        423 KFLLNLNKELQAERVIAITQQQWDNNIGDLAGISRLGEKVLSGKAYVDAFEEGKHIKADK 482

Qy        492 NITLDAKTGIIDISNSNGKKTQALHFTSPLLTAGTESRERLTNGKYSYINKLKFGRVKNW 551
               : ||:  ||||:|||    |||: | :|||| ||| |||:  ||| || ||    || :|
Db        483 LVQLDSANGIIDVSNSGKAKTQHILFRTPLLTPGTEHRERVQTGKYEYITKLNINRVDSW 542

Qy        552 QVTDGEASSKLDFSKVIQRV-------AETEGTDEIGLIVNAKAGNDDIFVGQGKMNIDG 604
              ::||| |||  | : |:||:            | |  :|    |:|::||| |   |||
Db        543 KITDGAASSTFDLTNVVQRIGIELDNAGNVTKTKETKIIAKLGEGDDNVFVGSGTTEIDG 602

Qy        605 GDGHDRVFYSKDGGFGNITVDGTSATEAGSYTVNRKVARGDIYHEVVKRQETKVGKRTET 664
              |:|:||| ||: | :| :|: |  | ||||||| |  | |  |||| ||     || | |
Db        603 GEGYDRVHYSR-GNYGALTIDATKETEQGSYTVNRFVETGKALHEVTSTHTALVGNREEK 661

Qy        665 IQYRDYELRKVGYGYQSTDNLKSVEEVIGSQFNDVFKGSKFNDIFHSGEGDDLLDGGAGD 724
              |:|| :    :    || : | |:|:|||:|:  ||:||||||||| |: |:| | :|| |:
Db        662 IEYR-HSNNQHHAGYYTKDTLKAVEEIIGTSHNDIFKGSKFNDAFNGGDGVDTIDGNDGN 720

Qy        725 DRLFGGKGNDRLSGDEGDDLLDGGSGDDVLNGGAGNDVYIFRKGDGNDTLYDGTGNDKLA 784
              |||||||:| | |  ||| :||| |:|:|:|| |:|::: |||||| : |  ||||:
Db        721 DRLFGGKGDDILDGGNGDDFIDGGKGNDLLHGGKGDDIFVHRKGDGNDIITDSDGNDKLS 780

Qy        785 FADANISDIMIERTKEGIIVKRNDHSGSINIPRWY----ITSNLQNYQSNKTDHKIEQLI 840
              |:|:|: |: |: | :::  |     : |  |:      :||:: | | |||::|
Db        781 FSDSNLKDLTFEKVKHNLVI-TNSKKEKVTIQNWFREADFAKEVPNYKATK-DEKIEEII 838

Qy        841 GKDGSYITSDQIDKILQDKKDGTVITSQELKKLADENKSQKLSASDIASSLNKLVGSMAL 900
              |::| ||| |:| ::  |     || ||| |:  |  | ::::|||:|| |:
Db        839 GQNGERITSKQVDDLI--AKGNGKITQDELSKVVDNYELLKHS-KNVTNSLDKLISSVSA 895

Qy        901 FGTANSVSSNALQPITQPTQGI 922
              | ::|   :  : |  : | :
Db        896 FTSSNDSRNVLVAPTSMLDQSL 917
```

RESULT 24
US-08-954-418-4
; Sequence 4, Application US/08954418
; Patent No. 6096320
; GENERAL INFORMATION:
; APPLICANT: POTTER, ANDREW
; APPLICANT: CAMPOS, MANUEL
; APPLICANT: HUGHES, HUW P.A.
; TITLE OF INVENTION: CYTOKINE-CYTOTOXIN GENE FUSIONS AND USES
; TITLE OF INVENTION: THEREOF
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: REED & ROBINS
; STREET: 635 BRYANT STREET
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/954,418
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/170,126
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/571,301
; ATTORNEY/AGENT INFORMATION:
; NAME: ROBINS, ROBERTA L.
; REGISTRATION NUMBER: 33,208
; REFERENCE/DOCKET NUMBER: 9000-0013.21
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 617-8999
; TELEFAX: (415) 327-3231
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1069 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-954-418-4

```
  Query Match             50.1%;  Score 2326.5;  DB 3;  Length 1069;
  Best Local Similarity   50.0%;  Pred. No. 1.9e-173;
  Matches  461;  Conservative  174;  Mismatches  262;  Indels   25;  Gaps   12;

Qy         15 LNSTKSGLKNLYLAIPKD--YDPQKGGTLNDFIKAADELGIARLAEEPNHTETAKKSVDT 72
              |:  |:| | : | ||::  || ::|  | | :|||:||||     || |:  ||: |: |
Db          7 LSFPKTGAKKIILYIPQNYQYDTEQGNGLQDLVKAAEELGIEVQREERNNIATAQTSLGT 66

Qy         73 VNQFLSLTQTGIAISATKLEKFLQKHSTNKLAKGLDSVENIDRKLGKASNVLSTLSSFLG 132
              :   : ||: || :|| :::| |||     |  : | | |:| :    ||  ||| : | ||
Db         67 IQTAIGLTERGIVLSAPQIDKLLQK---TKAGQALGSAESIVQNANKAKTVLSGIQSILG 123

Qy        133 TALAGIELDSLIKKGDAAPDALAKASIDLINEIIGNLSQSTQTIEAFSSQLAKLGSTISQ 192
              : |||::||   : : ::   ||||| ::| | :| |:: | :|:: |  |::: || :
Db        124 SVLAGMDLDEAL-QNNSNQHALAKAGLELTNSLIENIANSVKTLDEFGEQISQFGSKLQN 182

Qy        193 AKGFSNIGNKLQNL-NFSKTNLGLEIITGLLSGISAGFALADKNASTGKKVAAGFELSNQ 251
               ||    :|:||:|:    |  |||::|:||||| :|  ||||||| ||| |||||:||
Db        183 IKGLGTLGDKLKNIGGLDKAGLGLDVISGLLSGATAALVLADKNASTAKKVGAGFELANQ 242

Qy        252 VIGNVTKAISSYVLAQRVAAGLSTTGAVAALITSSIMLAISPLAFMNAADKFNHANALDE 311
              |:||:|||:|||:||||||||||:|| ||||| |:: ||||||||     |||||| :|:
Db        243 VVGNITKAVSSYILAQRVAAGLSSTGPVAALIASTVSLAISPLAFAGIADKFNHAKSLES 302

Qy        312 FAKQFRKFGYDGDHLLAEYQRGVGTIEASLTTISTALGAVSAGVSAAAVGSAVGTPIALL 371
              :|::|:| |||||:|||||||| |||:||:| |:||| |:: ||||| ||  : :|||||
Db        303 YAERFKKLGYDGDNLLAEYQRGTGTIDASVTAINTALAAIAGGVSAAAAGSVIASPIALL 362

Qy        372 VAGVTGLISGILEASKQAMFESVANRLQGKILEWEKQNGGQNYFDKGYDSRYAAYLANNL 431
              |:|:||:|| ||: |||||| |||::  ||:|||| | |:|||: |||:|| | | :|:
Db        363 VSGITGVISTILQYSKQAMFEHVANKIHNKIVEWEKNNHGKNYFENGYDARYLANLQDNM 422

Qy        432 KFLSELNKELEAERVIAITQQRWDNNIGELAGITKLGERIKSGKAYADAFEDGKKVEAGS 491
              |||  |||||:||||||||||:||||||:||||::|||:: |||||  ||||:|| ::|
Db        423 KFLLNLNKELQAERVIAITQQQWDNNIGDLAGISRLGEKVLSGKAYVDAFEEGKHIKADK 482
```

```
Qy        492 NITLDAKTGIIDISNSNGKKTQALHFTSPLLTAGTESRERLTNGKYSYINKLKFGRVKNW 551
Db        483 LVQLDSANGIIDVSNSGKAKTQHILFRTPLLTPGTEHRERVQTGKYEYITKLNINRVDSW 542

Qy        552 QVTDGEASSKLDFSKVIQRV-------AETEGTDEIGLIVNAKAGNDDIFVGQGKMNIDG 604
Db        543 KITDGAASSTFDLTNVVQRIGIELDNAGNVTKTKETKIIAKLGEGDDNVFVGSGTTEIDG 602

Qy        605 GDGHDRVFYSKDGGFGNITVDGTSATEAGSYTVNRKVARGDIYHEVVKRQETKVGKRTET 664
Db        603 GEGYDRVHYSR-GNYGALTIDATKETEQGSYTVNRFVETGKALHEVTSTHTALVGNREEK 661

Qy        665 IQYRDYELRKVGYGYQSTDNLKSVEEVIGSQFNDVFKGSKFNDIFHSGEGDDLLDGGAGD 724
Db        662 IEYR-HSNNQHHAGYYTKDTLKAVEEIIGTSHNDIFKGSKFNDAFNGGDGVDTIDGNDGN 720

Qy        725 DRLFGGKGNDRLSGDEGDDLLDGGSGDDVLNGGAGNDVYIFRKGDGNDTLYDGTGNDKLA 784
Db        721 DRLFGGKGDDILDGGNGDDFIDGGKGNDLLHGGKGDDIFVHRKGDGNDIITDSDGNDKLS 780

Qy        785 FADANISDIMIERTKEGIIVKRNDHSGSINIPRWY----ITSNLQNYQSNKTDHKIEQLI 840
Db        781 FSDSNLKDLTFEKVKHNLVI-TNSKKEKVTIQNWFREADFAKEVPNYKATK-DEKIEEII 838

Qy        841 GKDGSYITSDQIDKILQDKKDGTVITSQELKKLADENKSQKLSASDIASSLNKLVGSMAL 900
Db        839 GQNGERITSKQVDDLI--AKGNGKITQDELSKVVDNYELLKHS-KNVTNSLDKLISSVSA 895

Qy        901 FGTANSVSSNALQPITQPTQGI 922
Db        896 FTSSNDSRNVLVAPTSMLDQSL 917
```

RESULT 25
US-08-215-805A-80
; Sequence 80, Application US/08215805A
; Patent No. 5559008
; GENERAL INFORMATION:
; APPLICANT: Chang, Yung-Fu
; TITLE OF INVENTION: LEUKOTOXIN GENE FROM PASTEURELLA
; TITLE OF INVENTION: SUIS
; NUMBER OF SEQUENCES: 84
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon, Hargrave, Devans & Doyle
; STREET: Clinton Square, P.O. Box 1051
; CITY: Rochester
; STATE: New York
; COUNTRY: USA
; ZIP: 14603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/215,805A
; FILING DATE: 22-MAR-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Timian, Susan J.
; REGISTRATION NUMBER: 34,103
; REFERENCE/DOCKET NUMBER: 19603/61 (D-1329A)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (716) 263-1636
; TELEFAX: (716) 263-1600
; INFORMATION FOR SEQ ID NO: 80:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 934 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Pasteurella suis
; STRAIN: 5943
; IMMEDIATE SOURCE:
; LIBRARY: P. suis DNA in Bacteriophage lambda-dash
; CLONE: (Lambda)yfc33-37
US-08-215-805A-80

Query Match 49.4%; Score 2297.5; DB 1; Length 934;
Best Local Similarity 49.7%; Pred. No. 2.9e-171;
Matches 469; Conservative 172; Mismatches 254; Indels 49; Gaps 17;

```
Qy          1 MSNINV-IKSNIQAGLNST-----------KSGLKNLYLAIPKD--YDPQKGGTLNDFIK 46
Db          4 LANISTNLKNSLQSGLHKTGQSLNQAGQSLKAGAKKLILYIPKDYEYDSGRGNGLQDLVK 63

Qy         47 AADELGIARLAEEPNHTETAKKSVDTVNQFLSLTQTGIAISATKLEKFLQKHSTNKLAKG 106
Db         64 AAEDLGIEVQREERNGIATAQNSLSTIQNILGFSERGVVLSAPQLDKLLQKY---KISKA 120

Qy        107 LDSVENIDRKLGKASNVLSTLSSFLGTALAGIELDSLIKKGDAAPDALAKASIDLINEII 166
Db        121 PGSSENVAKNLGNAQTLLSGIQSILGSVMAGMDLDEILKNKGSELD-LAKAGLELTNSLI 179

Qy        167 GNLSQSTQTIEAFSSQLAKLGSTISQAKGFSNIGNKLQNLN-FSKTNLGLEIITGLLSGI 225
Db        180 ENIANSVQTLDTFSEQISQLGTKLQNVKGLGTLGDKLKNFSGFSKAGLGLEVISGLLSGA 239

Qy        226 SAGFALADKNASTGKKVAAGFELSNQVIGNVTKAISSYVLAQRVAAGLSTTGAVAALITS 285
Db        240 TAALVLADKNASTDRKVGAGFELANQVVGNITKAVSSYILAQRVAAGLSNTGPVSALIAS 299

Qy        286 SIMLAISPLAFMNAADKFNHANALDEFAKQFRKFGYDGDHLLAEYQRGVGTIEASLTTIS 345
Db        300 TVALAISPLAFAGIADKFNNAKALESYAERFKKLGYEGDSLLAEYQRGTGTIDASVTAVN 359

Qy        346 TALGAVSAGVSAAAVGSAVGTPIALLVAGVTGLISGILEASKQAMFESVANRLQGKILEW 405
Db        360 TALAAISGGVSAAAAGSLVGAPIALLVSGITGIISTILQYSKQAMFEHVANKIHDKIVDW 419

Qy        406 EKQNGGQNYFDKGYDSRYAAYLANNLKFLSELNKELEAERVIAITQQRWDNNIGELAGIT 465
Db        420 EKKHNGKNYFENGYDSRYLADLQDNMRQLQNLNKELQAERVIRITQQQWDNNIGNLAGIS 479

Qy        466 KLGERIKSGKAYADAFEDGKKVEAGSNITLDAKTGIIDISNSNGKKTQALHFTSPLLTAG 525
Db        480 RLGEKVMSGKAYADAFEEGKLIKADTFVQLDSATGVINTSKSDNVKTQHILFRTPLLTPG 539

Qy        526 TESRERLTNGKYSYINKLKFGRVKNWQVTDGEASSKLDFSKVIQRVA-------ETEGTD 578
Db        540 VENRERIQTGKYEYITKLNINRVDSWKITDGATNSTFDLTNVVQRIGIELDHADNVTKTK 599

Qy        579 EIGLIVNAKAGNDDIFVGQGKMNIDGGDGHDRVFYSKDGGFGNITVDGTSATEAGSYTVN 638
Db        600 ETKIIANLGDGNDDVFIGSGTTEVDGGNGLDRVHYSR-GDYGALTIDATNESVQGSYTVK 658

Qy        639 RKVARGDIYHEVVKRQETKVGKRTETIQYRDYELRKVGYGYQSTDNLKSVEEVIGSQFND 698
Db        659 RFVETGKALHEVTATQSVLVGSREEKIEYR-HSNNTQHAGYYTTDTLKSVEEIIGTSRND 717

Qy        699 VFKGSKFNDIFHSGEGDDLLDGGAGDDRLFGGKGNDRLSGDEGDDLLDGGSGDDVLNGGA 758
Db        718 IFKGSKFDDAFHGGDGVDNIDGNAGNDRLFGGKGFDIIDGGDGDDFIDGGQGDDILHGGK 777

Qy        759 GNDVYIFRKGDGNDTLYDGTGNDKLAFADANISDIMIERTKEGIIVKRNDHSGSINIPRW 818
Db        778 GNDILCTVKG-GNDSISDSGGNDRLSFADSNLKDLTFEKVNHHLMI-TNVKKEKVTIQNW 835

Qy        819 Y----ITSNLQNYQSNKTDHKIEQLIGKDGSYITSDQIDKILQDKKDGTVITSQELKKLA 874
Db        836 FREADYAKTVHNYQAT-ADEKIEEIIGRQGERITSKQIDELIEKGKG--KIDQSELERIA 892

Qy        875 DENKSQKLSASDIAS-SLNKLVGSMALFGTANSVSSNALQPITQ 917
```

```
Db        893 E--SSALLKESKFASNSLNKLVSSAAHL---------PLQTITE 925

RESULT 26
US-08-772-270A-8
; Sequence 8, Application US/08772270A
; Patent No. 6019984
;  GENERAL INFORMATION:
;    APPLICANT:  MacInnes, Janet
;    APPLICANT:  Ricciatti, Paul
;    APPLICANT:  Mallard, Bonnie
;    APPLICANT:  Rosendal, Soren
;    TITLE OF INVENTION:  NOVEL BACTERIAL PREPARATIONS, METHOD FOR
;    TITLE OF INVENTION:  PRODUCING SAME, AND THEIR USE AS VACCINES
;    NUMBER OF SEQUENCES:  14
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Bereskin & Parr
;      STREET:  40 King Street West
;      CITY:  Toronto
;      STATE:  Ontario
;      COUNTRY:  Canada
;      ZIP:  M5H 3Y2
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/772,270A
;      FILING DATE:  December 23, 1996
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Gravelle, Micheline
;      REGISTRATION NUMBER:  40,261
;      REFERENCE/DOCKET NUMBER:  6580-81
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (416) 364-7311
;      TELEFAX:  (416) 361-1398
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  956 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    ORIGINAL SOURCE:
;      ORGANISM:  Actinobacillus pleuropneumoniae
US-08-772-270A-8

  Query Match              48.1%;  Score 2236.5;  DB 3;  Length 956;
  Best Local Similarity    47.7%;  Pred. No. 1.8e-166;
  Matches  453;  Conservative  188;  Mismatches  255;  Indels   53;  Gaps   15;

Qy          1 MSNINV--IKSNIQAGL----------------------NSTKSGLKNLYLAIPKDYDPQ 36
Db          1 MSKITLSSLKSSLQQGLKNGKNKLNQAGTTLKNGLTQTGHSLQNGAKKLILYIPQGYDSG 60

Qy         37 KGGTLNDFIKAADELGIARLAEEPNHTETAKKSVDTVNQFLSLTQTGIAISATKLEKFLQ 96
Db         61 QGNGVQDLVKAANDLGIEVWREERSNLDIAKTSFDTTQKILGFTDRGIVLFAPQLDNLLK 120

Qy         97 KHSTNKLAKGLDSVENIDRKLGKASNVLSTLSSFLGTALAGIELDSLIKKGDAAPDALAK 156
Db        121 KNP--KIGNTLGSASSISQNIGKANTVLGGIQSILGSVLSGVNLNELLQNKDPNQLELAK 178

Qy        157 ASIDLINEIIGNLSQSTQTIEAFSSQLAKLGSTISQAKGFSNIGNKLQNL-NFSKTNLGL 215
Db        179 AGLELTNELVGNIASSVQTVDAFAEQISKLGSHLQNVKGLGGLSNKLQNLPDLGKASLGL 238

Qy        216 EIITGLLSGISAGFALADKNASTGKKVAAGFELSNQVIGNVTKAISSYVLAQRVAAGLST 275
Db        239 DIISGLLSGASAGLILADKEASTEKKAAAGVEFANQIIGNVTKAVSSYILAQRVASGLSS 298

Qy        276 TGAVAALITSSIMLAISPLAFMNAADKFNHANALDEFAKQFRKFGYDGDHLLAEYQRGVG 335
Db        299 TGPVAALIASTVALAVSPLSFLNVADKFKQADLIKSYSERFQKLGYDGDRLLADFHRETG 358

Qy        336 TIEASLTTISTALGAVSAGVSAAAVGSAVGTPIALLVAGVTGLISGILEASKQAMFESVA 395
Db        359 TIDASVTTINTALAAISGGVGAASAGSLVGAPVALLVAGVTGLITTILEYSKQAMFEHVA 418

Qy        396 NRLQGKILEWEKQNGGQNYFDKGYDSRYAAYLANNLKFLSELNKELEAERVIAITQQRWD 455
Db        419 NKVHDRIVEWEKKH-NKNYFEQGYDSRHLADLQDNMKFLINLNKELQAERVVAITQQRWD 477

Qy        456 NNIGELAGITKLGERIKSGKAYADAFEDGKKVEAGSNITLDAKTGIIDISNSNGKKTQAL 515
Db        478 NQIGDLAAISRRTDKISSGKAYVDAFEEGQHQSYDSSVQLDNKNGIINISNTN-RKTQSV 536

Qy        516 HFTSPLLTAGTESRERLTNGKYSYINKLKFGRVKNWQVTDGEASSKLDFSKVIQRVA--- 572
Db        537 LFRTPLLTPGEENRERIQEGKNSYITKLHIQRVDSWTVTDGDASSSVDFTNVVQRIAVKF 596

Qy        573 -------ETEGTDEIGLIVNAKAGNDDIFVGQGKMNIDGGDGHDRVFYSKDGGFGNITVD 625
Db        597 DDAGNIIESKDTK---IIANLGAGNDNVFVGSSTTVIDGGDGHDRVHYSR-GEYGALVID 652

Qy        626 GTSATEAGSYTVNRKVARGDIYHEVVKRQETKVGKRTETIQYRDYELRKVGYGYQSTDNL 685
Db        653 ATAETEKGSYSVKRYVGDSKALHETIATHQTNVGNREEKIEYR-REDDRFHTGYTVTDSL 711

Qy        686 KSVEEVIGSQFNDVFKGSKFNDIFHSGEGDDLLDGGAGDDRLFGGKGNDRLSGDEGDDLL 745
Db        712 KSVEEIIGSQFNDIFKGSQFDDVFHGGNGVDTIDGNDGDDHLFGGAGDDVIDGGNGNNFL 771

Qy        746 DGGSGDDVLNGGAGNDVYIFRKGDGNDTLYDGTGNDKLAFADANISDIMIERTKEGIIVK 805
Db        772 VGGTGNDIISGGKDNDIYVHKTGDGNDSITDSGGQDKLAFSDVNLKDLTFKKVDSSLEI- 830

Qy        806 RNDHSGSINIPRWY----ITSNLQNYQSNKTDHKIEQLIGKDGSYITSDQIDKILQDKKD 861
Db        831 INQKGEKVRIGNWFLEDDLASTVANYKAT-NDRKIEEIIGKGGERITSEQVDKLI--KEG 887

Qy        862 GTVITSQELKKLADENKSQKLSASDIASSLNKLVGSMALFGTANSVSSN 910
Db        888 NNQISAEALSKVVNDYNTSK-DRQNVSNSLAKLISSVGSFTSSSDFRNN 935

RESULT 27
US-09-062-126-8
; Sequence 8, Application US/09062126
; Patent No. 6500435
; GENERAL INFORMATION:
;  APPLICANT: Kamp, Elbarte Margriet
;  APPLICANT:  Smits, Marinus Adrianus
;  TITLE OF INVENTION: Recombinant Vaccine For  Prevention
;  TITLE OF INVENTION:  And/Or Treatment Of Pleuropneumonia Infections
;  FILE REFERENCE: 470-980537
;  CURRENT APPLICATION NUMBER: US/09/062,126
;  CURRENT FILING DATE:  1998-04-17
;  PRIOR APPLICATION NUMBER: 08/488,706
;  PRIOR FILING DATE: 1995-06-09
;  PRIOR APPLICATION NUMBER: 08/138,609
;  PRIOR FILING DATE: 1993-10-15
;  PRIOR APPLICATION NUMBER: 07/722,971
;  PRIOR FILING DATE: 1991-06-28
;  NUMBER OF SEQ ID NOS: 15
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 8
;   LENGTH: 956
;   TYPE: PRT
;   ORGANISM: Actinobacillus pleuropneumonia
US-09-062-126-8

  Query Match              48.1%;  Score 2236.5;  DB 4;  Length 956;
```

```
Best Local Similarity   47.7%;  Pred. No. 1.8e-166;
Matches  453;  Conservative  188;  Mismatches  255;  Indels   53;  Gaps   15;

Qy          1 MSNINV--IKSNIQAGL---------------------NSTKSGLKNLYLAIPKDYDPQ 36
Db          1 MSKITLSSLKSSLQQGLKNGKNKLNQAGTTLKNGLTQTGHSLQNGAKKLILYIPQGYDSG 60

Qy         37 KGGTLNDFIKAADELGIARLAEEPNHTETAKKSVDTVNQFLSLTQTGIAISATKLEKFLQ 96
Db         61 QGNGVQDLVKAANDLGIEVWREERSNLDIAKTSFDTTQKILGFTDRGIVLFAPQLDNLLK 120

Qy         97 KHSTNKLAKGLDSVENIDRKLGKASNVLSTLSSFLGTALAGIELDSLIKKGDAAPDALAK 156
Db        121 KNP--KIGNTLGSASSISQNIGKANTVLGGIQSILGSVLSGVNLNELLQNKDPNQLELAK 178

Qy        157 ASIDLINEIIGNLSQSTQTIEAFSSQLAKLGSTISQAKGFSNIGNKLQNL-NFSKTNLGL 215
Db        179 AGLELTNELVGNIASSVQTVDAFAEQISKLGSHLQNVKGLGGLSNKLQNLPDLGKASLGL 238

Qy        216 EIITGLLSGISAGFALADKNASTGKKVAAGFELSNQVIGNVTKAISSYVLAQRVAAGLST 275
Db        239 DIISGLLSGASAGLILADKEASTEKKAAAGVEFANQIIGNVTKAVSSYILAQRVASGLSS 298

Qy        276 TGAVAALITSSIMLAISPLAFMNAADKFNHANALDEFAKQFRKFGYDGDHLLAEYQRGVG 335
Db        299 TGPVAALIASTVALAVSPLSFLNVADKFKQADLIKSYSERFQKLGYDGDRLLADFHRETG 358

Qy        336 TIEASLTTISTALGAVSAGVSAAAVGSAVGTPIALLVAGVTGLISGILEASKQAMFESVA 395
Db        359 TIDASVTTINTALAAISGGVGAASAGSLVGAPVALLVAGVTGLITTILEYSKQAMFEHVA 418

Qy        396 NRLQGKILEWEKQNGGQNYFDKGYDSRYAAYLANNLKFLSELNKELEAERVIAITQQRWD 455
Db        419 NKVHDRIVEWEKKH-NKNYFEQGYDSRHLADLQDNMKFLINLNKELQAERVVAITQQRWD 477

Qy        456 NNIGELAGITKLGERIKSGKAYADAFEDGKKVEAGSNITLDAKTGIIDISNSNGKKTQAL 515
Db        478 NQIGDLAAISRRTDKISSGKAYVDAFEEGQHQSYDSSVQLDNKNGIINISNTN-RKTQSV 536

Qy        516 HFTSPLLTAGTESRERLTNGKYSYINKLKFGRVKNWQVTDGEASSKLDFSKVIQRVA--- 572
Db        537 LFRTPLLTPGEENRERIQEGKNSYITKLHIQRVDSWTVTDGDASSSVDFTNVVQRIAVKF 596

Qy        573 -------ETEGTDEIGLIVNAKAGNDDIFVGQGKMNIDGGDGHDRVFYSKDGGFGNITVD 625
Db        597 DDAGNIIESKDTK---IIANLGAGNDNVFVGSSTTVIDGGDGHDRVHYSR-GEYGALVID 652

Qy        626 GTSATEAGSYTVNRKVARGDIYHEVVKRQETKVGKRTETIQYRDYELRKVGYGYQSTDNL 685
Db        653 ATAETEKGSYSVKRYVGDSKALHETIATHQTNVGNREEKIEYR-REDDRFHTGYTVTDSL 711

Qy        686 KSVEEVIGSQFNDVFKGSKFNDIFHSGEGDDLLDGGAGDDRLFGGKGNDRLSGDEGDDLL 745
Db        712 KSVEEIIGSQFNDIFKGSQFDDVFHGGNGVDTIDGNDGDDHLFGGAGDDVIDGGNGNNFL 771

Qy        746 DGGSGDDVLNGGAGNDVYIFRKGDGNDTLYDGTGNDKLAFADANISDIMIERTKEGIIVK 805
Db        772 VGGTGNDIISGGKDNDIYVHKTGDGNDSITDSGGQDKLAFSDVNLKDLTFKKVDSSLEI- 830

Qy        806 RNDHSGSINIPRWY----ITSNLQNYQSNKTDHKIEQLIGKDGSYITSDQIDKILQDKKD 861
Db        831 INQKGEKVRIGNWFLEDDLASTVANYKAT-NDRKIEEIIGKGGERITSEQVDKLI--KEG 887

Qy        862 GTVITSQELKKLADENKSQKLSASDIASSLNKLVGSMALFGTANSVSSN 910
Db        888 NNQISAEALSKVVNDYNTSK-DRQNVSNSLAKLISSVGSFTSSSDFRNN 935
```

RESULT 28
US-09-062-126-10
; Sequence 10, Application US/09062126
; Patent No. 6500435
; GENERAL INFORMATION:
; APPLICANT: Kamp, Elbarte Margriet
; APPLICANT: Smits, Marinus Adrianus
; TITLE OF INVENTION: Recombinant Vaccine For Prevention
; TITLE OF INVENTION: And/Or Treatment Of Pleuropneumonia Infections
; FILE REFERENCE: 470-980537
; CURRENT APPLICATION NUMBER: US/09/062,126
; CURRENT FILING DATE: 1998-04-17
; PRIOR APPLICATION NUMBER: 08/488,706
; PRIOR FILING DATE: 1995-06-09
; PRIOR APPLICATION NUMBER: 08/138,609
; PRIOR FILING DATE: 1993-10-15
; PRIOR APPLICATION NUMBER: 07/722,971
; PRIOR FILING DATE: 1991-06-28
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 10
; LENGTH: 1053
; TYPE: PRT
; ORGANISM: Actinobacillus pleuropneumonia
US-09-062-126-10

```
  Query Match             41.4%;  Score 1926;  DB 4;  Length 1053;
  Best Local Similarity   42.1%;  Pred. No. 4.7e-142;
  Matches  420;  Conservative  168;  Mismatches  290;  Indels  120;  Gaps   19;

Qy          8 KSNIQAGLNSTKSGLK---------------------NLYLAIPKDYDPQKGGTLNDFI 45
Db         19 KRQAKKGYDVTKNGLQYGVSQAKLQALGAGKAVQKYGNKLVLVIPKEYDGSVGNGFFDLV 78

Qy         46 KAADELGIARLAEEPNHTETAKKSVDTVNQFLSLTQTGIAISATKLEKFLQKHSTNKLAK 105
Db         79 KAAEELGIQVKYVNRNELEVAHKTLGTADQFLGLTERGLTLFAPQLDQFLQKHSKISNVV 138

Qy        106 GLDSVENIDRKLGKASNVLSTLSSFLGTALAGIELDSLIKKGDAAPDALAKASIDLINEI 165
Db        139 GSSTGDAVS-KLAKSQTIISGIQSVLGTVLAGINLNEAIISGGSELE-LAEAGVSLASEL 196

Qy        166 IGNLSQSTQTIEAFSSQLAKLGSTISQAKGFSNIGNKLQNLN---FSKTNLGLEIITGLL 222
Db        197 LSNIAKGTTTIDAFTTQIQNFGKLVENAKGLGGVGRQLQNISGSALSKTGLGLDIISSLL 256

Qy        223 SGISAGFALADKNASTGKKVAAGFELSNQVIGNVTKAISSYVLAQRVAAGLSTTGAVAAL 282
Db        257 SGVTASFALANKNASTSTKVAAGFELSNQVIGGITKAVSSYILAQRLAAGLSSTGPAAAL 316

Qy        283 ITSSIMLAISPLAFMNAADKFNHANALDEFAKQFRKFGYDGDHLLAEYQRGVGTIEASLT 342
Db        317 IASSISLAISPLAFLRVADNFNRSKEIGEFAERFKKLGYDGDKLLSEFYHEAGTIDASIT 376

Qy        343 TISTALGAVSAGVSAAAVGSAVGTPIALLVAGVTGLISGILEASKQAMFESVANRLQGKI 402
Db        377 TISTALSAIAAGTAAASAGALVGAPITLLVTGITGLISGILEFSKQPMLDHVASKIGNKI 436

Qy        403 LEWEKQNGGQNYFDKGYDSRYAAYLANNLKFLSELNKELEAERVIAITQQRWDNNIGELA 462
Db        437 DEWEKKY-GKNYFENGYDARHKAFLEDSFSLLSSFNKQYETERAVLITQQRWDEYIGELA 495

Qy        463 GITKLGERIKSGKAYADAFEDGKKVEAG----SNITLDAKTGIIDISNSNGKKTQALHFT 518
Db        496 GITGKGDKLSSGKAYVDYFQEGKLLEKKPDDFSKVVFDPTKGEIDISNS--QTSTLLKFV 553

Qy        519 SPLLTAGTESRERLTNGKYSYINKLKFGRVKNWQVTDG--EASSKLDFSKVIQR------ 570
Db        554 TPLLTPGTESRERTQTGNYEYITKLVVKGKDKW-VVNGVKDKGAVYDYTNLIQHAHITSS 612

Qy        571 VAETEGTDEIGLIVNAKAGNDDIFVGQGKMNIDGGDGHDRVFYSKDGGFGNITVDGTSAT 630
Db        613 VARGEEYREVRLVSHLGNGNDKVFLVAGSAEIHAGEGHDVVYYDKT-DTGLLVIDGTKAT 671

Qy        631 EAGSYTVNRKVARG-DIYHEVVKRQETKVGKRTETIQYRDYELRKVG-YGYQSTDNLKSV 688
Db        672 EQGRYSVTRELSGATKILREVIKNQKSAVGKREETLEYRDYELTQSGNSNLKAHDELHSV 731
```

```
Qy        689 EEVIGSQFNDVFKGSKFNDIFHSGEGDDLLDGGAGDDRLFGGKGNDRLSGDEGDD----- 743
              ||:|||   | |||||| ||||  :|||||:|  ||| |:| |||| | || |:|
Db        732 EEIIGSNQRDEFKGSKFRDIFHGADGDDLLNGNDGDDILYGDKGNDELRGDNGNDQLYGG 791

Qy        744 -------------------------------------------------LLDGGSGDD 752
                                                               |||| |:|
Db        792 EGNDKLLGGNGNNYLSGGDGNDELQVLGKWFLMCFVAVKAMINFMVAQVLIYLDGGEGND 851

Qy        753 VLNGGAGNDVYIFRKGDGNDTLYD---GTGNDKLAFADANISDIMIERTKEGIIVKRNDH 809
                | || |:| |::    || |:||      :  |||  :| :    :::|:  : :::: |:
Db        852 YLEGGDGSDFYVYCSTSGNHTIYDQGKSSDLDKLYLSDFSFDRLLVEKVDDNLVLRSNES 911

Qy        810 S---GSINIPRWYITSNLQNYQSNKTDHKIEQLIGKDGSYITSDQIDKILQDKKDGTVIT 866
              |     |  :  |  |:          : || :||||:: :|    :|:: :      ::        :
Db        912 SHNNGVLTIKDWF-------KEGNKYNHKIEQIVDKNGRKLTAENLGTYFKNAPKADNLL 964

Qy        867 SQELKKLADENKSQKLSASDIASSLNKLVGSMALFGTA 904
              :     |:  |:|:|   :  | : :  |.:|:: :    || |
Db        965 NYATKE--DQNES---NLSSLKTELSKIITNAGNFGVA 997
```

RESULT 29
US-08-772-270A-11
; Sequence 11, Application US/08772270A
; Patent No. 6019984
; GENERAL INFORMATION:
; APPLICANT: MacInnes, Janet
; APPLICANT: Ricciatti, Paul
; APPLICANT: Mallard, Bonnie
; APPLICANT: Rosendal, Soren
; TITLE OF INVENTION: NOVEL BACTERIAL PREPARATIONS, METHOD FOR
; TITLE OF INVENTION: PRODUCING SAME, AND THEIR USE AS VACCINES
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Bereskin & Parr
; STREET: 40 King Street West
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5H 3Y2
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/772,270A
; FILING DATE: December 23, 1996
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Gravelle, Micheline
; REGISTRATION NUMBER: 40,261
; REFERENCE/DOCKET NUMBER: 6580-81
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 364-7311
; TELEFAX: (416) 361-1398
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1049 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Actinobacillus pleuropneumoniae
US-08-772-270A-11

```
  Query Match             41.4%;  Score 1924;  DB 3;  Length 1049;
  Best Local Similarity   43.0%;  Pred. No. 6.7e-142;
  Matches  428;  Conservative  152;  Mismatches  296;  Indels  120;  Gaps   21;

Qy          8 KSNIQAGLNSTKSGLK---------------------NLYLAIPKDYDPQKGGTLNDFI 45
              |  :: | : ||:||:                      | | |||:||   |    | :
Db         19 KRQVKKGYDVTKNGLQYGVSQAKLQALAAGKAVQKYGNKLVLVIPKEYDGSVGNGFFDLV 78

Qy         46 KAADELGIARLAEEPNHTETAKKSVDTVNQFLSLTQTGIAISATKLEKFLQKHSTNKLAK 105
              |||:||||     |  |  | ||:  | :|||  ||: |: :  |  :|::||||||
Db         79 KAAEELGIQVKYVNRNELEVAHKSLGTADQFLGLTERGLTLFAPQLDQFLQKHSKISNVV 138

Qy        106 GLDSVENIDRKLGKASNVLSTLSSFLGTALAGIELDSLIKKGDAAPDALAKASIDLINEI 165
              |  : : :  || |:  ::| :  | |||  ||||  |:   |    |    :  ||:|  :  | :|:
Db        139 GSSTGDAVS-KLAKSQTIISGIQSVLGTVLAGINLNEAIISGGSELE-LAEAGVSLASEL 196

Qy        166 IGNLSQSTQTIEAFSSQLAKLGSTISQAKGFSNIGNKLQNLN---FSKTNLGLEIITGLL 222
              : |::: | ||:||::|:   |        |||    :| :|||::     ||| |||:||: ||
Db        197 VSNIAKGTTTIDAFTTQIQNFGKLAENAKGLGGVGRQLQNISGSALSKTGLGLDIISSLL 256

Qy        223 SGIISAGFALADKNASTGKKVAAGFELSNQVIGNVTKAISSYVLAQRVAAGLSTTGAVAAL 282
              ||::  ||| :|||||  |||||||||||||| :|||:|||:||||:  |||||||  |||
Db        257 SGVTRSFALRNKNASTSTKVAAGFELSNQVIGGITKAVSSYILAQRLRAGLSTTGPAAAL 316

Qy        283 ITSSIMLAISPLAFMNAADKFNHANALDEFAKQFRKFGYDGDHLLAEYQRGVGTIEASLT 342
              | ||| ||||||||:  || ||  :  :  |||::|:| ||||| ||:|:       |||:||:|
Db        317 IASSISLAISPLAFLRVADNFNRSKEIGEFAERFKKLGYDGDKLLSEFYHEAGTIDASIT 376

Qy        343 TISTALGAVSAGVSAAAVGSAVGTPIALLVAGVTGLISGILEASKQAMFESVANRLQGKI 402
              |||||  |::|| :||: |:  || || ||| |:||||||||| ||| |  : ||::: ||
Db        377 TISTALSAIAAGTAAASAGALVGAPITLLVTGITGLISGILEFSKQPMLDHVASKIGNKI 436

Qy        403 LEWEKQNGGQNYFDKGYDSRYAAYLANNLKFLSELNKELEAERVIAITQQRWDNNIGELA 462
                ||||:  |:|||: |||:|: |:| ::     ||   ||: | || : ||||||  |||||
Db        437 DEWEKKY-GKNYFENGYDARHKAFLEDSFSLLSSFNKQYETERAVLITQQRWDEYIGELA 495

Qy        463 GITKLGERIKSGKAYADAFEDGKKVEAG----SNITLDAKTGIIDISNSNGKKTQALHFT 518
              |||   |:::  ||||  |  |::|| :                |  :       | ||||||      :    :    |
Db        496 GITGKGDKLSSGKAYVDYFQEGKLLEKKPDDFSKVVFDPTKGEIDISNS--QTSTLLKFV 553

Qy        519 SPLLTAGTESRERLTNGKYSYINKLKFGRVKNWQVTDG--EASSKLDFSKVIQR------ 570
              :||||  |||||||    ||| || ||              | | :|    :     :     |:: :||
Db        554 TPLLTPGTESRERTQTGKYEYITKLVVKGKDKW-VVNGVKDKGAVYDYTNLIQHAHISSS 612

Qy        571 VAETEGTDEIGLIVNAKAGNDDIFVGQGKMNIDGGDGHDRVFYSKDGGFGNITVDGTSAT 630
              ||   |     |: |: :     ||| :|:   |      |  |:||| |:|  |     |  : :||| ||
Db        613 VARGEEYREVRLVSHLGNGNDKVFLAAGSAEIHAGEGHDVVYYDKT-DTGLLVIDGTKAT 671

Qy        631 EAGSYTVNRKVARG-DIYHEVVKRQETKVGKRTETIQYRDYELRKVG-YGYQSTDNLKSV 688
              | | |:| |:::     |    ||:| |:   |||| ||::|||||| :  |      ::  |  |  ||
Db        672 EQGRYSVTRELSGATKILREVIKNQKYAVGKREETLEYRDYELTQSGNSNLKAHDELHSV 731

Qy        689 EEVIGSQFNDVFKGSKFNDIFHSGEGDDLLDGGAGDDRLFGGKGNDRLSGD--------- 739
              || |||    | |||||| ||||  :|||||:|  ||| |:| |||| | ||
Db        732 EE-IGSNQRDEFKGSKFRDIFHGADGDDLLNGNDGDDILYGDKGNDELRGDNGNDQLYGG 790

Qy        740 EGD----------------------------------------------DLLDGGSGDDV 753
              |||                                              |||||| |:|
Db        791 EGDDKLLGGNGNNYLSGGDGNDELQVLGNGFNVLRGGKGDDKLYGSSGSDLLDGGEGNDY 850

Qy        754 LNGGAGNDVYIFRKGDGNDTLYD---GTGNDKLAFADANISDIMIERTKEGIIVK--RND 808
              | || |:| |:: |     || |:||     : :|||  :| :  :|:::|    :  :   :       |
Db        851 LEGGDGSDFYVYRSTSGNHTIYDQGKASDSDKLYLSDLSFDNILVKRVNDNLEFRSNNNS 910

Qy        809 HSGSINIPRWYITSNLQNYQSNKTDHKIEQLIGKDGSYITSDQIDKILQDKKDGTVITSQ 868
              :|| :  | |:    |  |           |||||::|:|  :|:  :        |    |     |
Db        911 NSGVLTIKDWFKGGNSYN-------HKIEQIVDKNGRKLTAGNLGNNFHD----TQQASS 959

Qy        869 ELKKLADENKSQKLSASDIASSLNKLVGSMALFGTA 904
              || :  |      ||:  : :  | |:: :   || |
Db        960 LLKNVTQEQNESNLSS--LKTELGKIITNAGNFGVA 993
```

RESULT 30
PCT-US93-10500-2
; Sequence 2, Application PC/TUS9310500
; GENERAL INFORMATION:
; APPLICANT: Chang, Yung-Fu
; TITLE OF INVENTION: Recombinant Vaccine For Procine

```
;    TITLE OF INVENTION:  Pleuropneumoniae
;    NUMBER OF SEQUENCES:  2
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Alan S. Korman
;      STREET:  1600 Empire Tower
;      CITY:  Buffalo
;      STATE:  New York
;      COUNTRY:  U.S.A.
;      ZIP:  14202
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US93/10500
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/972,229
;      FILING DATE:  05-NOV-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Korman, Alan S.
;      REGISTRATION NUMBER:  33,932
;      REFERENCE/DOCKET NUMBER:  19603/00001
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  716-853-8104
;      TELEFAX:  716-853-8109
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1244 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    FRAGMENT TYPE:  N-terminal
;    ORIGINAL SOURCE:
;      ORGANISM:  Actinobacillus pleuropneumoniae
;      STRAIN:  Serotypes 2, 3, 4, 6 and 8
;      INDIVIDUAL ISOLATE:  Swine
;      CELL TYPE:  Gram negative bacterium
PCT-US93-10500-2

  Query Match             41.4%;  Score 1924;  DB 5;  Length 1244;
  Best Local Similarity   43.0%;  Pred. No. 8.7e-142;
  Matches  428;  Conservative  152;  Mismatches  296;  Indels  120;  Gaps   21;

Qy          8 KSNIQAGLNSTKSGLK---------------------NLYLAIPKDYDPQKGGTLNDFI 45
                |  :: | : ||:||:                      | | |||:||    |    | :
Db        192 KRQVKKGYDVTKNGLQYGVSQAKLQALAAGKAVQKYGNKLVLVIPKEYDGSVGNGFFDLV 251

Qy         46 KAADELGIARLAEEPNHTETAKKSVDTVNQFLSLTQTGIAISATKLEKFLQKHSTNKLAK 105
              |||:||||          |  |  | ||: | :|||  ||: |: : | :|::||||||
Db        252 KAAEELGIQVKYVNRNELEVAHKSLGTADQFLGLTERGLTLFAPQLDQFLQKHSKISNVV 311

Qy        106 GLDSVENIDRKLGKASNVLSTLSSFLGTALAGIELDSLIKKGDAAPDALAKASIDLINEI 165
              |  : : :  || |:  ::| : | ||| |||| |:  |   | :   : ||:| : | :|:
Db        312 GSSTGDAVS-KLAKSQTIISGIQSVLGTVLAGINLNEAIISGGSELE-LAEAGVSLASEL 369

Qy        166 IGNLSQSTQTIEAFSSQLAKLGSTISQAKGFSNIGNKLQNLN---FSKTNLGLEIITGLL 222
              : |::: | ||:||::|:   |       |||    :| :|||::    ||| |||:||: ||
Db        370 VSNIAKGTTTIDAFTTQIQNFGKLAENAKGLGGVGRQLQNISGSALSKTGLGLDIISSLL 429

Qy        223 SGISAGFALADKNASTGKKVAAGFELSNQVIGNVTKAISSYVLAQRVAAGLSTTGAVAAL 282
              ||::  ||| :|||||  |||||||||||| :|||:|||:||||: ||||||  |||
Db        430 SGVTRSFALRNKNASTSTKVAAGFELSNQVIGGITKAVSSYILAQRLRAGLSTTGPAAAL 489

Qy        283 ITSSIMLAISPLAFMNAADKFNHANALDEFAKQFRKFGYDGDHLLAEYQRGVGTIEASLT 342
              | ||| |||||||:  || ||  :   : |||::|:| ||||| ||:|:        |||:|:|
Db        490 IASSISLAISPLAFLRVADNFNRSKEIGEFAERFKKLGYDGDKLLSEFYHEAGTIDASIT 549

Qy        343 TISTALGAVSAGVSAAAVGSAVGTPIALLVAGVTGLISGILEASKQAMFESVANRLQGKI 402
              ||||||  |::|| :||: |: || || ||| |:||||||||| ||| |  : ||::: ||
Db        550 TISTALSAIAAGTAAASAGALVGAPITLLVTGITGLISGILEFSKQPMLDHVASKIGNKI 609

Qy        403 LEWEKQNGGQNYFDKGYDSRYAAYLANNLKFLSELNKELEAERVIAITQQRWDNNIGELA 462
              ||||:  |:|||: |||:|: |:| ::   ||  ||:`| || : |||||| |||||
Db        610 DEWEKKY-GKNYFENGYDARHKAFLEDSFSLLSSFNKQYETERAVLITQQRWDEYIGELA 668

Qy        463 GITKLGERIKSGKAYADAFEDGKKVEAG----SNITLDAKTGIIDISNSNGKKTQALHFT 518
              |||  |:::  |||||  | |::|| :|          |:  |   | ||||||  : :  | |
Db        669 GITGKGDKLSSGKAYVDYFQEGKLLEKKPDDFSKVVFDPTKGEIDISNS--QTSTLLKFV 726

Qy        519 SPLLTAGTESRERLTNGKYSYINKLKFGRVKNWQVTDG--EASSKLDFSKVIQR------ 570
              :||||  |||||||     ||| || ||           | | :|   :    :  |:: :||
Db        727 TPLLTPGTESRERTQTGKYEYITKLVVKGKDKW-VVNGVKDKGAVYDYTNLIQHAHISSS 785

Qy        571 VAETEGTDEIGLIVNAKAGNDDIFVGQGKMNIDGGDGHDRVFYSKDGGFGNITVDGTSAT 630
              ||  |     |: |: :     ||| :|:   |     |  |:||| |:| |     | : :||| ||
Db        786 VARGEEYREVRLVSHLGNGNDKVFLAAGSAEIHAGEGHDVVYYDKT-DTGLLVIDGTKAT 844

Qy        631 EAGSYTVNRKVARG-DIYHEVVKRQETKVGKRTETIQYRDYELRKVG-YGYQSTDNLKSV 688
              | | |:| |:::     |  ||:| |:  |||  ||::|||||  : |       :: | ||
Db        845 EQGRYSVTRELSGATKILREVIKNQKYAVGKREETLEYRDYELTQSGNSNLKAHDELHSV 904

Qy        689 EEVIGSQFNDVFKGSKFNDIFHSGEGDDLLDGGAGDDRLFGGKGNDRLSGD--------- 739
              || |||    | ||||||  ||||  :|||||:| ||| |:|  |||| | ||
Db        905 EE-IGSNQRDEFKGSKFRDIFHGADGDDLLNGNDGDDILYGDKGNDELRGDNGNDQLYGG 963

Qy        740 EGD-----------------------------------------------DLLDGGSGDDV 753
              |||                                              |||||| |:|
Db        964 EGDDKLLGGNGNNYLSGGDGNDELQVLGNGFNVLRGGKGDDKLYGSSGSDLLDGGEGNDY 1023

Qy        754 LNGGAGNDVYIFRKGDGNDTLYD---GTGNDKLAFADANISDIMIERTKEGIIVK--RND 808
              | || |:| |::|    || |:||       : :|||   :| :   :|:::|   : :  :   |
Db       1024 LEGGDGSDFYVYRSTSGNHTIYDQGKASDSDKLYLSDLSFDNILVKRVNDNLEFRSNNNS 1083

Qy        809 HSGSINIPRWYITSNLQNYQSNKTDHKIEQLIGKDGSYITSDQIDKILQDKKDGTVITSQ 868
              :|| :  |  |:   |   |        |||||:: |:|  :|: :       |    |   |
Db       1084 NSGVLTIKDWFKGGNSYN-------HKIEQIVDKNGRKLTAGNLGNNFHD----TQQASS 1132

Qy        869 ELKKLADENKSQKLSASDIASSLNKLVGSMALFGTA 904
               || :  |        ||:  : : | |:: :   |||
Db       1133 LLKNVTQEQNESNLSS--LKTELGKIITNAGNFGVA 1166

RESULT 31
US-08-772-270A-2
; Sequence 2, Application US/08772270A
; Patent No. 6019984
;  GENERAL INFORMATION:
;    APPLICANT:  MacInnes, Janet
;    APPLICANT:  Ricciatti, Paul
;    APPLICANT:  Mallard, Bonnie
;    APPLICANT:  Rosendal, Soren
;    TITLE OF INVENTION:  NOVEL BACTERIAL PREPARATIONS, METHOD FOR
;    TITLE OF INVENTION:  PRODUCING SAME, AND THEIR USE AS VACCINES
;    NUMBER OF SEQUENCES:  14
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Bereskin & Parr
;      STREET:  40 King Street West
;      CITY:  Toronto
;      STATE:  Ontario
;      COUNTRY:  Canada
;      ZIP:  M5H 3Y2
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/772,270A
;      FILING DATE:  December 23, 1996
;      CLASSIFICATION:  424
```

```
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Gravelle, Micheline
;      REGISTRATION NUMBER:  40,261
;      REFERENCE/DOCKET NUMBER:  6580-81
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (416) 364-7311
;      TELEFAX:  (416) 361-1398
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1022 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  circular
;    MOLECULE TYPE:  protein
;    ORIGINAL SOURCE:
;      ORGANISM:  Actinobacillus pleuropneumoniae
US-08-772-270A-2

  Query Match             37.9%;  Score 1763.5;  DB 3;  Length 1022;
  Best Local Similarity   40.0%;  Pred. No. 2.4e-129;
  Matches  400;  Conservative  181;  Mismatches  309;  Indels  109;  Gaps   26;

Qy          8 KSNIQAGLNSTKSGL-------KNLYLAIPKDYDPQKGGTLNDFIKAADELGIARLAEEP 60
Db         21 KSAAKSGAGALKNGLGQVKQAGQKLILYIPKDYQASTGSSLNDLVKAAEALGIEVHRSEK 80

Qy         61 NHTETAKKSVDTVNQFLSLTQTGIAISATKLEKFLQKHSTNKLAKGL-DSVENIDRKLGK 119
Db         81 NGTALAKELFGTTEKLLGFSERGIALFAPQFDKLLNKN--QKLSKSLGGSSEALGQRLNK 138

Qy        120 ASNVLSTLSSFLGTALAGIELDSLIKK----GDAAPDALAKASIDLINEIIGNLSQSTQT 175
Db        139 TQTALSALQSFLGTAIAGMDLDSLLRRRRNGEDVSGSELAKAGVDLAAQLVDNIASATGT 198

Qy        176 IEAFSSQLAKLGSTISQAKGFSNIGNKLQNL-NFSKTNLGLEIITGLLSGISAGFALADK 234
Db        199 VDAFAEQLGKLGNALSNTR-LSGLASKLNNLPDLSLAGPGFDAVSGILSVVSASFILSNK 257

Qy        235 NASTGKKVAAGFELSNQVIGNVTKAISSYVLAQRVAAGLSTTGAVAALITSSIMLAISPL 294
Db        258 DADAGTKAAAGIEISTKILGNIGKAVSQYIIAQRVAAGLSTTAATGGLIGSVVALAISPL 317

Qy        295 AFMNAADKFNHANALDEFAKQFRKFGYDGDHLLAEYQRGVGTIEASLTTISTALGAVSAG 354
Db        318 SFLNVADKFERAKQLEQYSERFKKFGYEGDSLLASFYRETGAIEAALTTINSVLSAASAG 377

Qy        355 VSAAAVGSAVGTPIALLVAGVTGLISGILEASKQAMFESVANRLQGKILEWEKQNGGQNY 414
Db        378 VGAAATGSLVGAPVAALVSAITGIISGILDASKQAIFERVATKLANKIDEWEKKH-GKNY 436

Qy        415 FDKGYDSRYAAYLANNLKFLSELNKELEAERVIAITQQRWDNNIGELAGITKLGERIKSG 474
Db        437 FENGYDARHSAFLEDTFELLSQYNKEYSVERVVAITQQRWDVNIGELAGITRKGSDTKSG 496

Qy        475 KAYADAFEDGKKVEAG----SNITLDAKTGIIDISNSNGKKTQALHFTSPLLTAGTESRE 530
Db        497 KAYVDFFEEGKLLEKEPDRFDKKVFDPLEGKIDLSSIN--KTTLLKFVTPVFTAGEEIRE 554

Qy        531 RLTNGKYSYINKLKFGRVKNWQVTDGEASSKL-DFSKVIQRVAETEGTDEIGLIVNAKAG 589
Db        555 RKQTGKYEYMTELFVKGKEKWVVTGVQSHNAIYDYTNLIQLAIDKKG-EKRQVTIESHLG 613

Qy        590 --NDDIFVGQGKMNIDGGDGHDRVFYSK-DGGFGNITVDGTSATEAGSYTVNRKV-ARGD 645
Db        614 EKNDRIYLSSGSSIVYAGNGHDVAYYDKTDTGY--LTFDGQSAQKAGEYIVTKELKADVK 671

Qy        646 IYHEVVKRQETKVGKRTETIQYRDYELR--KVGYGYQSTDNLKSVEEVIGSQFNDVFKGS 703
Db        672 VLKEVVKTQDISVGKRSEKLEYRDYELSPFELGNGIRAKDELHSVEEIIGSNRKDKFFGS 731

Qy        704 KFNDIFHSGEGD-----------------DLLDGGAGDDRLFGGKGNDRLSGDEGDDLL 745
Db        732 RFTDIFHGAKGDDEIYGNDGHDILYGDDGNDVIHGGDGNDHLVGGNGNDRLIGGKGNNFL 791

Qy        746 DGGSGDD----------VLNGGAGNDVY-------IFRKGDGNDTLYDGTGND------- 781
Db        792 NGGDGDDELQVFEGQYNVLLGGAGNDILYGSDGTNLFDGGVGNDKIYGGLGKDIYRYSKE 851

Qy        782 ------------------KLAFADANI----SDIMIERTKEGIIVKRNDHSG-SINIPR 817
Db        852 YGRHIIIEKGGDDDTLLLSDLSFKDVGFIRIGDDLLVNKRIGGTLYYHEDYNGNALTIKD 911

Qy        818 WYITSNLQNYQSNKTDHKIEQLIGKDGSYITSDQIDKILQDKKDGTVITSQELKKLADEN 877
Db        912 WF--KEGKEGQNN----KIEKIVDKDGAYVLSQYLTELTAPGRGINYFNGLEEKLYYGEG 965

Qy        878 KSQKLSASDIASSLNKLVGSMALF-GTANSVSSNALQPI 915
Db        966 YN---ALPQLRKDIEQIISSTGAFTGDHGKVSVGSGGPL 1001

RESULT 32
US-09-062-126-3
; Sequence 3, Application US/09062126
; Patent No. 6500435
; GENERAL INFORMATION:
;  APPLICANT: Kamp, Elbarte Margriet
;  APPLICANT:  Smits, Marinus Adrianus
;  TITLE OF INVENTION: Recombinant Vaccine For  Prevention
;  TITLE OF INVENTION:  And/Or Treatment Of Pleuropneumonia Infections
;  FILE REFERENCE: 470-980537
;  CURRENT APPLICATION NUMBER: US/09/062,126
;  CURRENT FILING DATE:  1998-04-17
;  PRIOR APPLICATION NUMBER: 08/488,706
;  PRIOR FILING DATE: 1995-06-09
;  PRIOR APPLICATION NUMBER: 08/138,609
;  PRIOR FILING DATE: 1993-10-15
;  PRIOR APPLICATION NUMBER: 07/722,971
;  PRIOR FILING DATE: 1991-06-28
;  NUMBER OF SEQ ID NOS: 15
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 3
;   LENGTH: 1013
;   TYPE: PRT
;   ORGANISM: Actinobacillus pleuropneumonia
US-09-062-126-3

  Query Match             37.7%;  Score 1751.5;  DB 4;  Length 1013;
  Best Local Similarity   40.1%;  Pred. No. 2.1e-128;
  Matches  401;  Conservative  180;  Mismatches  308;  Indels  111;  Gaps   27;

Qy          8 KSNIQAGLNSTKSGL-------KNLYLAIPKDYDPQKGGTLNDFIKAADELGIARLAEEP 60
Db         21 KSAAKSGAGALKNGLGQVKQAGQKLILYIPKDYQASTGSSLNDLVKAAEALGIEVHRSEK 80

Qy         61 NHTETAKKSVDTVNQFLSLTQTGIAISATKLEKFLQKHSTNKLAKGL-DSVENIDRKLGK 119
Db         81 NGTALAKELFGTTEKLLGFSERGIALFAPQFDKLLNKN--QKLSKSLGGSSEALGQRLNK 138

Qy        120 ASNVLSTLSSFLGTALAGIELDSLIKK----GDAAPDALAKASIDLINEIIGNLSQSTQT 175
Db        139 TQTALSALQSFLGTAIAGMDLDSLLRRRRNGEDDSGSELAKAGVDLAAQLVDNIASATGT 198

Qy        176 IEAFSSQLAKLGSTISQAKGFSNIGNKLQNL-NFSKTNLGLEIITGLLSGISAGFALADK 234
Db        199 VEAFAEQLGKLGNALSNTR-LSGLASKLNNLPDLSLAGPGFDAVSGILSVVSASFILSNK 257

Qy        235 NASTGKKVAAGFELSNQVIGNVTKAISSYVLAQRVAAGLSTTGAVAALITSSIMLAISPL 294
Db        258 DADAGTKAAAGIEISTKILGNIGKAVSQYIIAQRVAAGLSTTAATGGLIGSVVALAISPL 317

Qy        295 AFMNAADKFNHANALDEFAKQFRKFGYDGDHLLAEYQRGVGTIEASLTTISTALGAVSAG 354
Db        318 SFLIVADKFERAKQLEQYSERFKKFGYKGDSLLASFYRETGAIEAALTTINSVLSAASAG 377

Qy        355 VSAAAVGSAVGTPIALLVAGVTGLISGILEASKQAMFESVANRLQGKILEWEKQNGGQNY 414
```

```
Db        378 VGAAATGSLVGAPVAALVSAITGIISGILDASKQAIFERVATKLANKIDEWEKKH-GKNY 436

Qy        415 FDKGYDSRYAAYLANNLKFLSELNKELEAERVIAITQQRWDNNIGELAGITKLGERIKSG 474

Db        437 FENGYDARHSAFLEDTFELLSQYNKEYSVERVVAITQQRWDVNIGELAGITRKGSDTKSG 496

Qy        475 KAYADAFEDGKKVEAGSNITLDAKT-----GIIDISNSNGKKTQALHFTSPLLTAGTESR 529

Db        497 KAYVDFFEEGKLLEKEPD-RFDKKVKDPLEGKIDLSSIN--KTTLLKFVTPVFTAGEEIR 553

Qy        530 ERLTNGKYSYINKLKFGRVKNWQVTDGEASSKL-DFSKVIQRVAETEGTDEIGLIVNAKA 588

Db        554 ERKQTGKYEYMTELFVKGKEKWVVTGVQSHNAIYDYTNLIQLAIDKKG-EKRQVTIESHL 612

Qy        589 G--NDDIFVGQGKMNIDGGDGHDRVFYSK-DGGFGNITVDGTSATEAGSYTVNRKV-ARG 644

Db        613 GEKNDRIYLSSGSSIVYAGNGHDVAYYDKTDTGY--LTFDGQSAQKAGEYIVTKELKADV 670

Qy        645 DIYHEVVKRQETKVGKRTETIQYRDYELR--KVGYGYQSTDNLKSVEEVIGSQFNDVFKG 702

Db        671 KVLKEVVKTQDISVGKRSEKLEYRDYELSPFELGNGIRAKDELHSVEEIIGSNRKDKFFG 730

Qy        703 SKFNDIFHSGEGD-----------------DLLDGGAGDDRLFGGKGNDRLSGDEGDDL 744

Db        731 SRFTDIFHGAKGDDEIYGNDGHDILYGDDGNDVIHGGDGNDHLVGGNGNDRLIGGKGNNF 790

Qy        745 LDGGSGDD---------VLNGGAGNDVY-------IFRKGDGNDTLYDGTGND------ 781

Db        791 LNGGDGDDELQVFEGQYNVLLGGAGNDILYGSDGTNLFDGGVGNDKIYGGLGKDIYRYSK 850

Qy        782 -------------------KLAFADANI----SDIMIERTKEGIIVKRNDHSG-SINIP 816

Db        851 EYGRHIIIEKGGDDDTLLLSDLSFKDVGFIRIGDDLLVNKRIGGTLYYHEDYNGNALTIK 910

Qy        817 RWYITSNLQNYQSNKTDHKIEQLIGKDGSYITSDQIDKILQDKKDGTVITSQELKKLADE 876

Db        911 DWF--KEGKEGQNN----KIEKIVDKDGAYVLSQYLTELTAPGRGINYFNGLEEKLYYGE 964

Qy        877 NKSQKLSASDIASSLNKLVGSMALF-GTANSVSSNALQPI 915

Db        965 GYN---ALPQLRKDIEQIISSTGAFTGDHGKVSVGSGGPL 1001
```

RESULT 33
US-08-258-188-2
; Sequence 2, Application US/08258188
; Patent No. 5475098
; GENERAL INFORMATION:
; APPLICANT: HALL, Robert H.
; APPLICANT: XU, Jian Guo
; TITLE OF INVENTION: A NEW AND DISTINCTIVE DNA SEQUENCE OF E.
; TITLE OF INVENTION: coli O157:H7 AND ITS USE FOR THE RAPID, SENSITIVE AND
; TITLE OF INVENTION: SPECIFIC DETECTION OF O157:H7 AND OTHER ENTEROHEMORRHAGIC
; TITLE OF INVENTION: E. coli
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Khourie and Crew
; STREET: Steuart Street Tower, One Market Plaza
; CITY: San Francisco
; STATE: California
; COUNTRY: US
; ZIP: 94105-1493
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/258,188
; FILING DATE: 14-JUN-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Garrett-Wackowski, Eugenia
; REGISTRATION NUMBER: 37,330
; REFERENCE/DOCKET NUMBER: 15280206,DHHSE135940
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 543-9600
; TELEFAX: (415) 543-5043
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 758 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-258-188-2

```
  Query Match             29.6%;  Score 1376.5;  DB 1;  Length 758;
  Best Local Similarity   40.5%;  Pred. No. 3.2e-99;
  Matches  304;  Conservative  129;  Mismatches  225;  Indels   93;  Gaps   16;

Qy        229 FALADKNASTGKKVAAGFELSNQVIGNVTKAISSYVLAQRVAAGLSTTGAVAALITSSIM 288

Db          2 FILGNSDAHTGTKAAAGIELTTQVLGNVGKAVSQYILAQRMAQGLSTTAASAGLITSAVM 61

Qy        289 LAISPLAFMNAADKFNHANALDEFAKQFRKFGYDGDHLLAEYQRGVGTIEASLTTISTAL 348

Db         62 LAISPLSFLAAADKFERAKQLESYSERFKKLNYEGDALLAAFHKETGAIDAALTTINTVL 121

Qy        349 GAVSAGVSAAAVGSAVGTPIALLVAGVTGLISGILEASKQAMFESVANRLQGKILEWEKQ 408

Db        122 SSVSAGVSAASSASLIGAPISMLVSALTGTISGILEASKQAMFEHVAEKFAARINEWEKE 181

Qy        409 NGGQNYFDKGYDSRYAAYLANNLKFLSELNKELEAERVIAITQQRWDNNIGELAGITKLG 468

Db        182 H-GKNYFENGYDARHAAFLEDSLSLLADFSRQHAVERAVAITQQHWDEKIGELAGITRNA 240

Qy        469 ERIKSGKAYADAFEDGKKVEAG----SNITLDAKTGIIDISNSNGKKTQALHFTSPLLTA 524

Db        241 DRSQSGKAYINYLENGGLLEAQPKEFTQQVFDPQKGTIDL--STGNVSSVLTFITPTFTP 298

Qy        525 GTESRERLTNGKYSYINKLKFGRVKNWQVTDGEASSK--LDFSKVIQRVAETEGTDEIGL 582

Db        299 GEEVRERKQSGKYEYMTSLIVNGKDTWSV-KGIKNHKGVYDYSKLIQFVEKNNKHYQARI 357

Qy        583 IVNAKAGNDDIFVGQGKMNIDGGDGHDRVFYSKDGGFGNITVDGTSATEAGSYTVNRKVA 642

Db        358 ISELGDKDDVVYSGAGSSEVFAGEGYDTVSYNKT-DVGKLTIDATGASKPGEYIVSKNM- 415

Qy        643 RGD--IYHEVVKRQETKVGKRTETIQYRDYELRKVGYGYQSTDNLKSVEEVIGSQFNDVF 700

Db        416 YGDVKVLQEVVKEQEVSVGKRTEKIQYRDFEFRTGGIPYDVIDNLHSVEELIGGKHDDEF 475

Qy        701 KGSKFNDIFHSG-----------------EGDDLLDGGAGDDRLFGGKGNDRLSGDEGD 742

Db        476 KGGKFNDIFHGADGNDYIEGNYGNDRLYGDDGDDYISGGQGDDQLFGGSGNDKLSGGDGN 535

Qy        743 -----------------------------------DLLDGGSGDDVLNGGAGNDVYIFR 766

Db        536 NYLTGGSGNDELQAHGAYNILSGGTGDDKLYGGGGIDLLDGGEGNDYLNGGFGNDIYVYG 595

Qy        767 KGDGNDTLYD-GTGNDKLAFADANISDIMIERTKEGIIVKR-----------NDHSGSIN 814

Db        596 QNYGHHTIADEGGKGDRLHLSDISFDDIAFKRVGNDLIMNKAINGVLSFNESNDVNG-IT 654

Qy        815 IPRWYITSNLQNYQSNKTDHKIEQLIGKDGSYITSDQIDKILQDKKDGTVITSQELKKLA 874

Db        655 FKNWFAKD-----ASGADNHLVEVITDKDGREI---KVDKIPHNNNE----RSGYIKASN 702

Qy        875 DENKSQKLSASDIASSLNKLVGSMALFGTAN 905

Db        703 IASEKNMVNITSVANDINKIISSVSGFDSGD 733
```

RESULT 34
US-08-526-813-2

```
; Sequence 2, Application US/08526813
; Patent No. 5756293
;  GENERAL INFORMATION:
;    APPLICANT:  Hall, Robert H.
;    APPLICANT:  Xu, Jian Guo
;    TITLE OF INVENTION:  A New and Distinctive DNA Sequence of E.
; Patent No. 5756293
;    TITLE OF INVENTION:  coli O157:H7 and its Use for the Rapid, Sensitive and
;    TITLE OF INVENTION:  Specific Detection of O157:H7 and Other Enterohemorrhagic
;    TITLE OF INVENTION:  E. coli
;    NUMBER OF SEQUENCES:  29
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend and Crew LLP
;      STREET:  Two Embarcadero Center, Eighth Floor
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94111-3834
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/526,813
;      FILING DATE:  11-SEP-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/258,188
;      FILING DATE:  14-JUN-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Garrett-Wackowski, Eugenia
;      REGISTRATION NUMBER:  37,330
;      REFERENCE/DOCKET NUMBER:  15280-206-1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 576-0200
;      TELEFAX:  (415) 576-0300
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  758 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-526-813-2

  Query Match             29.6%;  Score 1376.5;  DB 1;  Length 758;
  Best Local Similarity   40.5%;  Pred. No. 3.2e-99;
  Matches  304;  Conservative  129;  Mismatches  225;  Indels   93;  Gaps   16;

Qy        229 FALADKNASTGKKVAAGFELSNQVIGNVTKAISSYVLAQRVAAGLSTTGAVAALITSSIM 288

Db          2 FILGNSDAHTGTKAAAGIELTTQVLGNVGKAVSQYILAQRMAQGLSTTAASAGLITSAVM 61

Qy        289 LAISPLAFMNAADKFNHANALDEFAKQFRKFGYDGDHLLAEYQRGVGTIEASLTTISTAL 348

Db         62 LAISPLSFLAAADKFERAKQLESYSERFKKLNYEGDALLAAFHKETGAIDAALTTINTVL 121

Qy        349 GAVSAGVSAAAVGSAVGTPIALLVAGVTGLISGILEASKQAMFESVANRLQGKILEWEKQ 408

Db        122 SSVSAGVSAASSASLIGAPISMLVSALTGTISGILEASKQAMFEHVAEKFAARINEWEKE 181

Qy        409 NGGQNYFDKGYDSRYAAYLANNLKFLSELNKELEAERVIAITQQRWDNNIGELAGITKLG 468

Db        182 H-GKNYFENGYDARHAAFLEDSLSLLADFSRQHAVERAVAITQQHWDEKIGELAGITRNA 240

Qy        469 ERIKSGKAYADAFEDGKKVEAG----SNITLDAKTGIIDISNSNGKKTQALHFTSPLLTA 524

Db        241 DRSQSGKAYINYLENGGLLEAQPKEFTQQVFDPQKGTIDL--STGNVSSVLTFITPTFTP 298

Qy        525 GTESRERLTNGKYSYINKLKFGRVKNWQVTDGEASSK--LDFSKVIQRVAETEGTDEIGL 582

Db        299 GEEVRERKQSGKYEYMTSLIVNGKDTWSV-KGIKNHKGVYDYSKLIQFVEKNNKHYQARI 357

Qy        583 IVNAKAGNDDIFVGQGKMNIDGGDGHDRVFYSKDGGFGNITVDGTSATEAGSYTVNRKVA 642

Db        358 ISELGDKDDVVYSGAGSSEVFAGEGYDTVSYNKT-DVGKLTIDATGASKPGEYIVSKNM- 415

Qy        643 RGD--IYHEVVKRQETKVGKRTETIQYRDYELRKVGYGYQSTDNLKSVEEVIGSQFNDVF 700

Db        416 YGDVKVLQEVVKEQEVSVGKRTEKIQYRDFEFRTGGIPYDVIDNLHSVEELIGGKHDDEF 475

Qy        701 KGSKFNDIFHSG-----------------EGDDLLDGGAGDDRLFGGKGNDRLSGDEGD 742

Db        476 KGGKFNDIFHGADGNDYIEGNYGNDRLYGDDGDDYISGGQGDDQLFGGSGNDKLSGGDGN 535

Qy        743 ----------------------------------DLLDGGSGDDVLNGGAGNDVYIFR 766

Db        536 NYLTGGSGNDELQAHGAYNILSGGTGDDKLYGGGGIDLLDGGEGNDYLNGGFGNDIYVYG 595

Qy        767 KGDGNDTLYD-GTGNDKLAFADANISDIMIERTKEGIIVKR-----------NDHSGSIN 814

Db        596 QNYGHHTIADEGGKGDRLHLSDISFDDIAFKRVGNDLIMNKAINGVLSFNESNDVNG-IT 654

Qy        815 IPRWYITSNLQNYQSNKTDHKIEQLIGKDGSYITSDQIDKILQDKKDGTVITSQELKKLA 874

Db        655 FKNWFAKD-----ASGADNHLVEVITDKDGREI---KVDKIPHNNNE----RSGYIKASN 702

Qy        875 DENKSQKLSASDIASSLNKLVGSMALFGTAN 905

Db        703 IASEKNMVNITSVANDINKIISSVSGFDSGD 733


RESULT 35
PCT-US95-08554-2
; Sequence 2, Application PC/TUS9508554
;  GENERAL INFORMATION:
;    APPLICANT:  HALL, Robert H.
;    APPLICANT:  XU, Jian Guo
;    TITLE OF INVENTION:  A NEW AND DISTINCTIVE DNA SEQUENCE OF E.
;    TITLE OF INVENTION:  coli O157:H7 AND ITS USE FOR THE RAPID, SENSITIVE AND
;    TITLE OF INVENTION:  SPECIFIC DETECTION OF O157:H7 AND OTHER ENTEROHEMORRHAGIC
;    TITLE OF INVENTION:  E. coli
;    NUMBER OF SEQUENCES:  29
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend Khourie and Crew
;      STREET:  Steuart Street Tower, One Market Plaza
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  US
;      ZIP:  94105-1493
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/08554
;      FILING DATE:  14-JUN-1994
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Garrett-Wackowski, Eugenia
;      REGISTRATION NUMBER:  37,330
;      REFERENCE/DOCKET NUMBER:  15280206,DHHSE135940
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 543-9600
;      TELEFAX:  (415) 543-5043
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  758 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
PCT-US95-08554-2

  Query Match             29.6%;  Score 1376.5;  DB 5;  Length 758;
  Best Local Similarity   40.5%;  Pred. No. 3.2e-99;
```

```
  Matches  304;  Conservative  129;  Mismatches  225;  Indels   93;  Gaps   16;

Qy        229 FALADKNASTGKKVAAGFELSNQVIGNVTKAISSYVLAQRVAAGLSTTGAVAALITSSIM 288
Db          2 FILGNSDAHTGTKAAAGIELTTQVLGNVGKAVSQYILAQRMAQGLSTTAASAGLITSAVM 61

Qy        289 LAISPLAFMNAADKFNHANALDEFAKQFRKFGYDGDHLLAEYQRGVGTIEASLTTISTAL 348
Db         62 LAISPLSFLAAADKFERAKQLESYSERFKKLNYEGDALLAAFHKETGAIDAALTTINTVL 121

Qy        349 GAVSAGVSAAAVGSAVGTPIALLVAGVTGLISGILEASKQAMFESVANRLQGKILEWEKQ 408
Db        122 SSVSAGVSAASSASLIGAPISMLVSALTGTISGILEASKQAMFEHVAEKFAARINEWEKE 181

Qy        409 NGGQNYFDKGYDSRYAAYLANNLKFLSELNKELEAERVIAITQQRWDNNIGELAGITKLG 468
Db        182 H-GKNYFENGYDARHAAFLEDSLSLLADFSRQHAVERAVAITQQHWDEKIGELAGITRNA 240

Qy        469 ERIKSGKAYADAFEDGKKVEAG----SNITLDAKTGIIDISNSNGKKTQALHFTSPLLTA 524
Db        241 DRSQSGKAYINYLENGGLLEAQPKEFTQQVFDPQKGTIDL--STGNVSSVLTFITPTFTP 298

Qy        525 GTESRERLTNGKYSYINKLKFGRVKNWQVTDGEASSK--LDFSKVIQRVAETEGTDEIGL 582
Db        299 GEEVRERKQSGKYEYMTSLIVNGKDTWSV-KGIKNHKGVYDYSKLIQFVEKNNKHYQARI 357

Qy        583 IVNAKAGNDDIFVGQGKMNIDGGDGHDRVFYSKDGGFGNITVDGTSATEAGSYTVNRKVA 642
Db        358 ISELGDKDDVVYSGAGSSEVFAGEGYDTVSYNKT-DVGKLTIDATGASKPGEYIVSKNM- 415

Qy        643 RGD--IYHEVVKRQETKVGKRTETIQYRDYELRKVGYGYQSTDNLKSVEEVIGSQFNDVF 700
Db        416 YGDVKVLQEVVKEQEVSVGKRTEKIQYRDFEFRTGGIPYDVIDNLHSVEELIGGKHDDEF 475

Qy        701 KGSKFNDIFHSG-----------------EGDDLLDGGAGDDRLFGGKGNDRLSGDEGD 742
Db        476 KGGKFNDIFHGADGNDYIEGNYGNDRLYGDDGDDYISGGQGDDQLFGGSGNDKLSGGDGN 535

Qy        743 -----------------------------------DLLDGGSGDDVLNGGAGNDVYIFR 766
Db        536 NYLTGGSGNDELQAHGAYNILSGGTGDDKLYGGGGIDLLDGGEGNDYLNGGFGNDIYVYG 595

Qy        767 KGDGNDTLYD-GTGNDKLAFADANISDIMIERTKEGIIVKR----------NDHSGSIN 814
Db        596 QNYGHHTIADEGGKGDRLHLSDISFDDIAFKRVGNDLIMNKAINGVLSFNESNDVNG-IT 654

Qy        815 IPRWYITSNLQNYQSNKTDHKIEQLIGKDGSYITSDQIDKILQDKKDGTVITSQELKKLA 874
Db        655 FKNWFAKD-----ASGADNHLVEVITDKDGREI---KVDKIPHNNNE----RSGYIKASN 702

Qy        875 DENKSQKLSASDIASSLNKLVGSMALFGTAN 905
Db        703 IASEKNMVNITSVANDINKIISSVSGFDSGD 733
```

RESULT 36
US-08-387-156-10
; Sequence 10, Application US/08387156
; Patent No. 5723129
; GENERAL INFORMATION:
; APPLICANT: POTTER, ANDREW A.
; APPLICANT: REDMOND, MARK J.
; APPLICANT: HUGHES, HUW P.A.
; TITLE OF INVENTION: GnRH-LEUKOTOXIN CHIMERAS
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: REED & ROBINS
; STREET: 635 BRYANT STREET
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/387,156
; FILING DATE: 10-FEB-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/960,932
; FILING DATE: 14-OCT-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/779,171
; FILING DATE: 16-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: ROBINS, ROBERTA L.
; REGISTRATION NUMBER: 33,208
; REFERENCE/DOCKET NUMBER: 9001-0016.21
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 617-8999
; TELEFAX: (415) 327-3231
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 544 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-387-156-10

```
  Query Match             19.5%;  Score 904.5;  DB 1;  Length 544;
  Best Local Similarity   38.7%;  Pred. No. 1.9e-62;
  Matches  220;  Conservative   98;  Mismatches  169;  Indels   81;  Gaps   13;

Qy         15 LNSTKSGLKNLYLAIPKD--YDPQKGGTLNDFIKAADELGIARLAEEPNHTETAKKSVDT 72
Db          7 LSFPKTGAKKIILYIPQNYQYDTEQGNGLQDLVKAAEELGIEVQREERNNIATAQTSLGT 66

Qy         73 VNQFLSLTQTGIAISATKLEKFLQKHSTNKLAKGLDSVENIDRKLGKASNVLSTLSSFLG 132
Db         67 IQTAIGLTERGIVLSAPQIDKLLQK---TKAGQALGSAESIVQNANKAKTVLSGIQSILG 123

Qy        133 TALAGIELDSLIKKGDAAPDALAKASIDLINEIIGNLSQSTQTIEAFSSQLAKLGSTISQ 192
Db        124 SVLAGMDLDEAL-QNNSNQHALAKAGLELTNSLIENIANSVKTLDEFGEQISQFGSKLQN 182

Qy        193 AKGFSNIGNKLQNL-NFSKTNLGLEIITGLLSGISAGFALADKNASTGKKVAAGFELSNQ 251
Db        183 IKGLGTLGDKLKNIGGLDKAGLGLDVISGLLSGATAALVLADKNASTAKKVGAGFELANQ 242

Qy        252 VIGNVTKAISSYVLAQRVAAGLSTTGAVAALITSSIMLAISPLAFMNAADKFNHANALDE 311
Db        243 VVGNITKAVSSYILAQRVAAGLSSTGPVAALIASTVSLAISPLAFAGIADKFNHAKSLES 302

Qy        312 FAKQFRKFGYDGDHLLAEYQRGVGTIEASLTTISTALGAVSAGVSAAAVGSAVGTPIALL 371
Db        303 YAERFKKLGYDGDNLLAEYQRGTGTIDASVTAINTALAAIAGGVSAAAAN---------- 352

Qy        372 VAGVTGLISGILEASKQAMFESVANRLQGKILEWEKQNGGQNYFDKGYDSRYAAYLANNL 431
Db        353 --------------LKDLTFEKVKHNLVITNSKKEKVT-IQNWF---------------- 381

Qy        432 KFLSELNKELEAERVIAITQQRWDNNIGELAGITKLGERIKSGKAYADAFEDGKKVEAGS 491
Db        382 -------READFAKEVPNYKATKDEKIEEIIG--QNGERITSKQV------DDLIAKGNG 426

Qy        492 NITLDAKTGIID----------ISNSNGKKTQALHFTSPLLTAGTESRERL---TNGKYS 538
Db        427 KITQDELSKVVDNYELLKHSKNVTNSLDKLISSV----SAFTSSNDSRNVLVAPTSMLDQ 482

Qy        539 YINKLKFGR-VKNWQVTDGEASSKLDFS 565
Db        483 SLSSLQFARGSQHWSYGLRPGSGSQDWS 510
```

```
RESULT 37
US-08-694-865-10
; Sequence 10, Application US/08694865
; Patent No. 5837268
;  GENERAL INFORMATION:
;    APPLICANT:  POTTER, ANDREW A.
;    APPLICANT:  MANNS, JOHN G.
;    TITLE OF INVENTION:  GnRH-LEUKOTOXIN CHIMERAS
;    NUMBER OF SEQUENCES:  34
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  REED & ROBINS LLP
;      STREET:  285 HAMILTON AVENUE, SUITE 200
;      CITY:  PALO ALTO
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94301
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/694,865
;      FILING DATE:  09-AUG-1996
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  MCCRACKEN, THOMAS P.
;      REGISTRATION NUMBER:  38,548
;      REFERENCE/DOCKET NUMBER:  9001-0016.22
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415)327-3400
;      TELEFAX:  (415)327-3231
;  INFORMATION FOR SEQ ID NO:  10:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  544 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-694-865-10

  Query Match             19.5%;  Score 904.5;  DB 2;  Length 544;
  Best Local Similarity   38.7%;  Pred. No. 1.9e-62;
  Matches  220;  Conservative   98;  Mismatches  169;  Indels   81;  Gaps   13;

Qy         15 LNSTKSGLKNLYLAIPKD--YDPQKGGTLNDFIKAADELGIARLAEEPNHTETAKKSVDT 72
              |:  |:| | : | ||::  || ::|  | | :|||:||||    || |:  ||: |: |
Db          7 LSFPKTGAKKIILYIPQNYQYDTEQGNGLQDLVKAAEELGIEVQREERNNIATAQTSLGT 66

Qy         73 VNQFLSLTQTGIAISATKLEKFLQKHSTNKLAKGLDSVENIDRKLGKASNVLSTLSSFLG 132
              :    : ||: || :|| :::| |||       |  :  | | |:|  :    ||  ||| : | ||
Db         67 IQTAIGLTERGIVLSAPQIDKLLQK---TKAGQALGSAESIVQNANKAKTVLSGIQSILG 123

Qy        133 TALAGIELDSLIKKGDAAPDALAKASIDLINEIIGNLSQSTQTIEAFSSQLAKLGSTISQ 192
              : |||::||   : : ::       ||||| ::| | :| |:: | :|:: |   |::: || :
Db        124 SVLAGMDLDEAL-QNNSNQHALAKAGLELTNSLIENIANSVKTLDEFGEQISQFGSKLQN 182

Qy        193 AKGFSNIGNKLQNL-NFSKTNLGLEIITGLLSGISAGFALADKNASTGKKVAAGFELSNQ 251
               ||    :|:||:|:      |  |||::|:||||| :|   ||||||| ||| |||||:||
Db        183 IKGLGTLGDKLKNIGGLDKAGLGLDVISGLLSGATAALVLADKNASTAKKVGAGFELANQ 242

Qy        252 VIGNVTKAISSYVLAQRVAAGLSTTGAVAALITSSIMLAISPLAFMNAADKFNHANALDE 311
              |:||:|||:|||:||||||||||:|| ||||| |:: ||||||| ||||||| :|:
Db        243 VVGNITKAVSSYILAQRVAAGLSSTGPVAALIASTVSLAISPLAFAGIADKFNHAKSLES 302

Qy        312 FAKQFRKFGYDGDHLLAEYQRGVGTIEASLTTISTALGAVSAGVSAAAVGSAVGTPIALL 371
              :|::|:| ||||:||||||| |||:||:| |:||| |:: ||||||
Db        303 YAERFKKLGYDGDNLLAEYQRGTGTIDASVTAINTALAAIAGGVSAAAAN---------- 352

Qy        372 VAGVTGLISGILEASKQAMFESVANRLQGKILEWEKQNGGQNYFDKGYDSRYAAYLANNL 431
                                |  || | : |        : ||      ||:|
Db        353 --------------LKDLTFEKVKHNLVITNSKKEKVT-IQNWF---------------- 381

Qy        432 KFLSELNKELEAERVIAITQQRWDNNIGELAGITKLGERIKSGKAYADAFEDGKKVEAGS 491
                         :| :  : :    :     |  | | |: |  : |||| | :           |     :
Db        382 -------READFAKEVPNYKATKDEKIEEIIG--QNGERITSKQV------DDLIAKGNG 426

Qy        492 NITLDAKTGIID----------ISNSNGKKTQALHFTSPLLTAGTESRERL---TNGKYS 538
               || |   : ::|           ::||  |   ::          |:  :||  |    |:
Db        427 KITQDELSKVVDNYELLKHSKNVTNSLDKLISSV----SAFTSSNDSRNVLVAPTSMLDQ 482

Qy        539 YINKLKFGR-VKNWQVTDGEASSKLDFS 565
               :: |:| |   ::|        |   |:|
Db        483 SLSSLQFARGSQHWSYGLRPGSGSQDWS 510

RESULT 38
US-08-878-748-10
; Sequence 10, Application US/08878748
; Patent No. 5969126
;  GENERAL INFORMATION:
;    APPLICANT:  POTTER, ANDREW A.
;    APPLICANT:  REDMOND, MARK J.
;    APPLICANT:  HUGHES, HUW P.A.
;    TITLE OF INVENTION:  GnRH-LEUKOTOXIN CHIMERAS
;    NUMBER OF SEQUENCES:  28
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  REED & ROBINS
;      STREET:  635 BRYANT STREET
;      CITY:  PALO ALTO
;      STATE:  CALIFORNIA
;      COUNTRY:  UNITED STATES OF AMERICA
;      ZIP:  94301
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/878,748
;      FILING DATE:  19-JUN-1997
;      CLASSIFICATION:  536
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/387,156
;      FILING DATE:  10-FEB-1995
;      APPLICATION NUMBER:  US 07/960,932
;      FILING DATE:  14-OCT-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/779,171
;      FILING DATE:  16-OCT-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  ROBINS, ROBERTA L.
;      REGISTRATION NUMBER:  33,208
;      REFERENCE/DOCKET NUMBER:  9001-0016.21
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 617-8999
;      TELEFAX:  (415) 327-3231
;  INFORMATION FOR SEQ ID NO:  10:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  544 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-878-748-10

  Query Match             19.5%;  Score 904.5;  DB 2;  Length 544;
  Best Local Similarity   38.7%;  Pred. No. 1.9e-62;
  Matches  220;  Conservative   98;  Mismatches  169;  Indels   81;  Gaps   13;

Qy         15 LNSTKSGLKNLYLAIPKD--YDPQKGGTLNDFIKAADELGIARLAEEPNHTETAKKSVDT 72
              |:  |:| | : | ||::  || ::|  | | :|||:||||    || |:  ||: |: |
Db          7 LSFPKTGAKKIILYIPQNYQYDTEQGNGLQDLVKAAEELGIEVQREERNNIATAQTSLGT 66

Qy         73 VNQFLSLTQTGIAISATKLEKFLQKHSTNKLAKGLDSVENIDRKLGKASNVLSTLSSFLG 132
              :    : ||: || :|| :::| |||       |  :  | | |:|  :    ||  ||| : | ||
Db         67 IQTAIGLTERGIVLSAPQIDKLLQK---TKAGQALGSAESIVQNANKAKTVLSGIQSILG 123
```

```
Qy        133 TALAGIELDSLIKKGDAAPDALAKASIDLINEIIGNLSQSTQTIEAFSSQLAKLGSTISQ 192
              : |||::||  : : ::    ||||| ::| | :| |:: | :|:: |  |::: || :
Db        124 SVLAGMDLDEAL-QNNSNQHALAKAGLELTNSLIENIANSVKTLDEFGEQISQFGSKLQN 182

Qy        193 AKGFSNIGNKLQNL-NFSKTNLGLEIITGLLSGISAGFALADKNASTGKKVAAGFELSNQ 251
              ||    :|:||:|:     |  |||::|:||||| :|    ||||||| ||| |||||:||
Db        183 IKGLGTLGDKLKNIGGLDKAGLGLDVISGLLSGATAALVLADKNASTAKKVGAGFELANQ 242

Qy        252 VIGNVTKAISSYVLAQRVAAGLSTTGAVAALITSSIMLAISPLAFMNAADKFNHANALDE 311
              |:||:|||:|||:||||||||||:|| ||||| |:: |||||||    |||||| :|:
Db        243 VVGNITKAVSSYILAQRVAAGLSSTGPVAALIASTVSLAISPLAFAGIADKFNHAKSLES 302

Qy        312 FAKQFRKFGYDGDHLLAEYQRGVGTIEASLTTISTALGAVSAGVSAAAVGSAVGTPIALL 371
              :|::|:| |||||:||||||||| |||:||:| |:||| |:: ||||||
Db        303 YAERFKKLGYDGDNLLAEYQRGTGTIDASVTAINTALAAIAGGVSAAAAN---------- 352

Qy        372 VAGVTGLISGILEASKQAMFESVANRLQGKILEWEKQNGGQNYFDKGYDSRYAAYLANNL 431
                            |   || | : |      : ||     ||:|
Db        353 --------------LKDLTFEKVKHNLVITNSKKEKVT-IQNWF---------------- 381

Qy        432 KFLSELNKELEAERVIAITQQRWDNNIGELAGITKLGERIKSGKAYADAFEDGKKVEAGS 491
                     :| :  :  :   :   |  | |: |  : |||| | :        |     :
Db        382 -------READFAKEVPNYKATKDEKIEEIIG--QNGERITSKQV------DDLIAKGNG 426

Qy        492 NITLDAKTGIID----------ISNSNGKKTQALHFTSPLLTAGTESRERL---TNGKYS 538
               || |  : ::|          ::|| |  ::        |: :|| |  |:
Db        427 KITQDELSKVVDNYELLKHSKNVTNSLDKLISSV----SAFTSSNDSRNVLVAPTSMLDQ 482

Qy        539 YINKLKFGR-VKNWQVTDGEASSKLDFS 565
               :: |:| |  ::|       |  |:|
Db        483 SLSSLQFARGSQHWSYGLRPGSGSQDWS 510
```

RESULT 39
US-09-124-491-10

```
; Sequence 10, Application US/09124491
; Patent No. 6022960
;  GENERAL INFORMATION:
;    APPLICANT:  POTTER, ANDREW A.
;    APPLICANT:  MANNS, JOHN G.
;    TITLE OF INVENTION:  GnRH-LEUKOTOXIN CHIMERAS
;    NUMBER OF SEQUENCES:  34
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  REED & ROBINS LLP
;      STREET:  285 HAMILTON AVENUE, SUITE 200
;      CITY:  PALO ALTO
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94301
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/124,491
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/694,865
;      FILING DATE:  09-AUG-1996
;      APPLICATION NUMBER:  US 08/387,156
;      FILING DATE:  10-FEB-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/960,932
;      FILING DATE:  14-OCT-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/779,171
;      FILING DATE:  16-OCT-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  MCCRACKEN, THOMAS P.
;      REGISTRATION NUMBER:  38,548
;      REFERENCE/DOCKET NUMBER:  9001-0016.22
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415)327-3400
;      TELEFAX:  (415)327-3231
;  INFORMATION FOR SEQ ID NO:  10:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  544 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-124-491-10

  Query Match             19.5%;  Score 904.5;  DB 3;  Length 544;
  Best Local Similarity   38.7%;  Pred. No. 1.9e-62;
  Matches  220;  Conservative   98;  Mismatches  169;  Indels   81;  Gaps   13;

Qy         15 LNSTKSGLKNLYLAIPKD--YDPQKGGTLNDFIKAADELGIARLAEEPNHTETAKKSVDT 72
              |: :| | : | ||::  || ::|  | | :|||:||||      || |:  ||: |: |
Db          7 LSFPKTGAKKIILYIPQNYQYDTEQGNGLQDLVKAAEELGIEVQREERNNIATAQTSLGT 66

Qy         73 VNQFLSLTQTGIAISATKLEKFLQKHSTNKLAKGLDSVENIDRKLGKASNVLSTLSSFLG 132
               :  : ||: || :|| :::| |||      |  : | | |:| :    ||  ||| : | ||
Db         67 IQTAIGLTERGIVLSAPQIDKLLQK---TKAGQALGSAESIVQNANKAKTVLSGIQSILG 123

Qy        133 TALAGIELDSLIKKGDAAPDALAKASIDLINEIIGNLSQSTQTIEAFSSQLAKLGSTISQ 192
              : |||::||  : : ::    ||||| ::| | :| |:: | :|:: |  |::: || :
Db        124 SVLAGMDLDEAL-QNNSNQHALAKAGLELTNSLIENIANSVKTLDEFGEQISQFGSKLQN 182

Qy        193 AKGFSNIGNKLQNL-NFSKTNLGLEIITGLLSGISAGFALADKNASTGKKVAAGFELSNQ 251
              ||    :|:||:|:     |  |||::|:||||| :|    ||||||| ||| |||||:||
Db        183 IKGLGTLGDKLKNIGGLDKAGLGLDVISGLLSGATAALVLADKNASTAKKVGAGFELANQ 242

Qy        252 VIGNVTKAISSYVLAQRVAAGLSTTGAVAALITSSIMLAISPLAFMNAADKFNHANALDE 311
              |:||:|||:|||:||||||||||:|| ||||| |:: |||||||    |||||| :|:
Db        243 VVGNITKAVSSYILAQRVAAGLSSTGPVAALIASTVSLAISPLAFAGIADKFNHAKSLES 302

Qy        312 FAKQFRKFGYDGDHLLAEYQRGVGTIEASLTTISTALGAVSAGVSAAAVGSAVGTPIALL 371
              :|::|:| |||||:||||||||| |||:||:| |:||| |:: ||||||
Db        303 YAERFKKLGYDGDNLLAEYQRGTGTIDASVTAINTALAAIAGGVSAAAAN---------- 352

Qy        372 VAGVTGLISGILEASKQAMFESVANRLQGKILEWEKQNGGQNYFDKGYDSRYAAYLANNL 431
                            |   || | : |      : ||     ||:|
Db        353 --------------LKDLTFEKVKHNLVITNSKKEKVT-IQNWF---------------- 381

Qy        432 KFLSELNKELEAERVIAITQQRWDNNIGELAGITKLGERIKSGKAYADAFEDGKKVEAGS 491
                     :| :  :  :   :   |  | |: |  : |||| | :        |     :
Db        382 -------READFAKEVPNYKATKDEKIEEIIG--QNGERITSKQV------DDLIAKGNG 426

Qy        492 NITLDAKTGIID----------ISNSNGKKTQALHFTSPLLTAGTESRERL---TNGKYS 538
               || |  : ::|          ::|| |  ::        |: :|| |  |:
Db        427 KITQDELSKVVDNYELLKHSKNVTNSLDKLISSV----SAFTSSNDSRNVLVAPTSMLDQ 482

Qy        539 YINKLKFGR-VKNWQVTDGEASSKLDFS 565
               :: |:| |  ::|       |  |:|
Db        483 SLSSLQFARGSQHWSYGLRPGSGSQDWS 510
```

RESULT 40
US-09-383-912-10

```
; Sequence 10, Application US/09383912
; Patent No. 6521746
;  GENERAL INFORMATION:
;    APPLICANT:  POTTER, ANDREW A.
;    APPLICANT:  MANNS, JOHN G.
;    TITLE OF INVENTION:  GnRH-LEUKOTOXIN CHIMERAS
;    NUMBER OF SEQUENCES:  34
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  REED & ROBINS LLP
;      STREET:  285 HAMILTON AVENUE, SUITE 200
;      CITY:  PALO ALTO
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94301
```

```
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/383,912
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/694,865
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  MCCRACKEN, THOMAS P.
;      REGISTRATION NUMBER:  38,548
;      REFERENCE/DOCKET NUMBER:  9001-0016.22
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415)327-3400
;      TELEFAX:  (415)327-3231
;  INFORMATION FOR SEQ ID NO:  10:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  544 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-383-912-10

  Query Match             19.5%;  Score 904.5;  DB 4;  Length 544;
  Best Local Similarity   38.7%;  Pred. No. 1.9e-62;
  Matches  220;  Conservative   98;  Mismatches  169;  Indels   81;  Gaps   13;

Qy         15 LNSTKSGLKNLYLAIPKD--YDPQKGGTLNDFIKAADELGIARLAEEPNHTETAKKSVDT 72

Db          7 LSFPKTGAKKIILYIPQNYQYDTEQGNGLQDLVKAAEELGIEVQREERNNIATAQTSLGT 66

Qy         73 VNQFLSLTQTGIAISATKLEKFLQKHSTNKLAKGLDSVENIDRKLGKASNVLSTLSSFLG 132

Db         67 IQTAIGLTERGIVLSAPQIDKLLQK---TKAGQALGSAESIVQNANKAKTVLSGIQSILG 123

Qy        133 TALAGIELDSLIKKGDAAPDALAKASIDLINEIIGNLSQSTQTIEAFSSQLAKLGSTISQ 192

Db        124 SVLAGMDLDEAL-QNNSNQHALAKAGLELTNSLIENIANSVKTLDEFGEQISQFGSKLQN 182

Qy        193 AKGFSNIGNKLQNL-NFSKTNLGLEIITGLLSGISAGFALADKNASTGKKVAAGFELSNQ 251

Db        183 IKGLGTLGDKLKNIGGLDKAGLGLDVISGLLSGATAALVLADKNASTAKKVGAGFELANQ 242

Qy        252 VIGNVTKAISSYVLAQRVAAGLSTTGAVAALITSSIMLAISPLAFMNAADKFNHANALDE 311

Db        243 VVGNITKAVSSYILAQRVAAGLSSTGPVAALIASTVSLAISPLAFAGIADKFNHAKSLES 302

Qy        312 FAKQFRKFGYDGDHLLAEYQRGVGTIEASLTTISTALGAVSAGVSAAAVGSAVGTPIALL 371

Db        303 YAERFKKLGYDGDNLLAEYQRGTGTIDASVTAINTALAAIAGGVSAAAAN---------- 352

Qy        372 VAGVTGLISGILEASKQAMFESVANRLQGKILEWEKQNGGQNYFDKGYDSRYAAYLANNL 431

Db        353 --------------LKDLTFEKVKHNLVITNSKKEKVT-IQNWF--------------- 381

Qy        432 KFLSELNKELEAERVIAITQQRWDNNIGELAGITKLGERIKSGKAYADAFEDGKKVEAGS 491

Db        382 -------READFAKEVPNYKATKDEKIEEIIG--QNGERITSKQV------DDLIAKGNG 426

Qy        492 NITLDAKTGIID----------ISNSNGKKTQALHFTSPLLTAGTESRERL---TNGKYS 538

Db        427 KITQDELSKVVDNYELLKHSKNVTNSLDKLISSV----SAFTSSNDSRNVLVAPTSMLDQ 482

Qy        539 YINKLKFGR-VKNWQVTDGEASSKLDFS 565

Db        483 SLSSLQFARGSQHWSYGLRPGSGSQDWS 510

Search completed: February 17, 2004, 10:11:33
```

Job time : 29 secs

```
OM protein - protein search, using sw model

Run on:         February 17, 2004, 10:12:24 ; Search time 39 Seconds
                                             (without alignments)
                                             4976.856 Million cell updates/sec

Title:          US-10-069-799-5
Perfect score:  4647
Sequence:       1 MSNINVIKSNIQAGLNSTKS..........SSNALQPITQPTQGILAPSV 927

Scoring table:  BLOSUM62
                Gapop 10.0 , Gapext 0.5

Searched:       801455 seqs, 209382283 residues

Total number of hits satisfying chosen parameters:      801455

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       Published_Applications_AA:*
                1:  /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
                2:  /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
                3:  /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
                4:  /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
                5:  /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
                6:  /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
                7:  /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
                8:  /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
                9:  /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
                10:  /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
                11:  /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
                12:  /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
                13:  /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
                14:  /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
                15:  /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
                16:  /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
                17:  /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
                18:  /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	4642	99.9	927	11	US-09-884-696-2	Sequence 2, Appli
2	2325	50.0	953	11	US-09-884-696-3	Sequence 3, Appli
3	2236.5	48.1	956	11	US-09-884-696-4	Sequence 4, Appli
4	1845	39.7	1023	11	US-09-884-696-5	Sequence 5, Appli
5	901	19.4	695	11	US-09-305-924-13	Sequence 13, Appl
6	408.5	8.8	1115	15	US-10-193-950A-2	Sequence 2, Appli
7	295.5	6.4	1017	12	US-10-369-493-20166	Sequence 20166, A
8	263.5	5.7	4327	12	US-10-369-493-10178	Sequence 10178, A
9	232	5.0	1778	12	US-10-238-075-749	Sequence 749, App
10	230.5	5.0	499	12	US-10-369-493-20175	Sequence 20175, A
11	226.5	4.9	2122	10	US-09-813-214A-9	Sequence 9, Appli
12	219	4.7	2285	10	US-09-932-183A-2	Sequence 2, Appli
13	216	4.6	2039	15	US-10-192-584-7	Sequence 7, Appli
14	211.5	4.6	2042	15	US-10-192-584-6	Sequence 6, Appli
15	202	4.3	2353	10	US-09-797-862-33	Sequence 33, Appl
16	200	4.3	239	12	US-10-369-493-19032	Sequence 19032, A
17	199	4.3	3241	9	US-09-841-786-1	Sequence 1, Appli
18	197	4.2	1833	12	US-10-175-275-4	Sequence 4, Appli
19	197	4.2	1833	12	US-10-175-282-4	Sequence 4, Appli
20	197	4.2	1992	12	US-10-175-275-3	Sequence 3, Appli
21	197	4.2	1992	12	US-10-175-282-3	Sequence 3, Appli
22	196	4.2	2086	9	US-09-815-242-5639	Sequence 5639, Ap
23	196	4.2	5795	9	US-09-815-242-12610	Sequence 12610, A
24	191	4.1	262	12	US-10-369-493-19792	Sequence 19792, A
25	188.5	4.1	1098	10	US-09-797-862-32	Sequence 32, Appl
26	188	4.0	1626	15	US-10-185-990-11	Sequence 11, Appl
27	186.5	4.0	2659	12	US-10-311-879-28	Sequence 28, Appl
28	186.5	4.0	3169	15	US-10-114-170-257	Sequence 257, App
29	186.5	4.0	6281	9	US-09-815-242-12996	Sequence 12996, A
30	180.5	3.9	1136	12	US-10-369-493-19046	Sequence 19046, A
31	180	3.9	210	12	US-10-369-493-20095	Sequence 20095, A
32	180	3.9	273	12	US-10-369-493-20096	Sequence 20096, A
33	179	3.9	773	9	US-09-841-786-6	Sequence 6, Appli
34	179	3.9	992	12	US-10-193-764-57	Sequence 57, Appl
35	179	3.9	998	12	US-10-193-764-55	Sequence 55, Appl
36	177.5	3.8	2434	9	US-09-815-242-5835	Sequence 5835, Ap
37	176.5	3.8	1741	10	US-09-971-536-68	Sequence 68, Appl
38	175.5	3.8	1228	10	US-09-117-447-2	Sequence 2, Appli
39	175	3.8	2893	12	US-09-882-227-522	Sequence 522, App
40	174.5	3.8	1536	12	US-10-193-764-63	Sequence 63, Appl
41	174	3.7	2124	12	US-10-369-493-11841	Sequence 11841, A
42	172	3.7	637	12	US-10-369-493-14403	Sequence 14403, A
43	172	3.7	2834	12	US-10-085-959-252	Sequence 252, App
44	170.5	3.7	1536	14	US-10-092-880-2	Sequence 2, Appli
45	169	3.6	1073	12	US-10-193-764-45	Sequence 45, Appl

ALIGNMENTS

```
RESULT 1
US-09-884-696-2
; Sequence 2, Application US/09884696
; Publication No. US20030035809A1
; GENERAL INFORMATION:
;  APPLICANT: GEORGE, LISLE W
;  APPLICANT: ANGELOS, JOHN A
;  APPLICANT:  HESS, JOHN F
;  TITLE OF INVENTION: MORAXELLA BOVIS CYTOTOXIN, CYTOTOXIN GENE, ANTIBODIES
;  TITLE OF INVENTION:  AND VACCINES FOR PREVENTION AND TREATMENT OF MORAXELLA
;  TITLE OF INVENTION:  BOVIS INFECTIONS
;  FILE REFERENCE: 481.06
;  CURRENT APPLICATION NUMBER: US/09/884,696
;  CURRENT FILING DATE:  2001-06-19
;  NUMBER OF SEQ ID NOS: 41
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
;   LENGTH: 927
;   TYPE: PRT
;   ORGANISM: Moraxella bovis
US-09-884-696-2

  Query Match             99.9%;  Score 4642;  DB 11;  Length 927;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches  926;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 MSNINVIKSNIQAGLNSTKSGLKNLYLAIPKDYDPQKGGTLNDFIKAADELGIARLAEEP 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MSNINVIKSNIQAGLNSTKSGLKNLYLAIPKDYDPQKGGTLNDFIKAADELGIARLAEEP 60

Qy         61 NHTETAKKSVDTVNQFLSLTQTGIAISATKLEKFLQKHSTNKLAKGLDSVENIDRKLGKA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 NHTETAKKSVDTVNQFLSLTQTGIAISATKLEKFLQKHSTNKLAKGLDSVENIDRKLGKA 120

Qy        121 SNVLSTLSSFLGTALAGIELDSLIKKGDAAPDALAKASIDLINEIIGNLSQSTQTIEAFS 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 SNVLSTLSSFLGTALAGIELDSLIKKGDAAPDALAKASIDLINEIIGNLSQSTQTIEAFS 180
```

```
Qy        181 SQLAKLGSTISQAKGFSNIGNKLQNLNFSKTNLGLEIITGLLSGISAGFALADKNASTGK 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 SQLAKLGSTISQAKGFSNIGNKLQNLNFSKTNLGLEIITGLLSGISAGFALADKNASTGK 240

Qy        241 KVAAGFELSNQVIGNVTKAISSYVLAQRVAAGLSTTGAVAALITSSIMLAISPLAFMNAA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 KVAAGFELSNQVIGNVTKAISSYVLAQRVAAGLSTTGAVAALITSSIMLAISPLAFMNAA 300

Qy        301 DKFNHANALDEFAKQFRKFGYDGDHLLAEYQRGVGTIEASLTTISTALGAVSAGVSAAAV 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 DKFNHANALDEFAKQFRKFGYDGDHLLAEYQRGVGTIEASLTTISTALGAVSAGVSAAAV 360

Qy        361 GSAVGTPIALLVAGVTGLISGILEASKQAMFESVANRLQGKILEWEKQNGGQNYFDKGYD 420
              ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GSAVGAPIALLVAGVTGLISGILEASKQAMFESVANRLQGKILEWEKQNGGQNYFDKGYD 420

Qy        421 SRYAAYLANNLKFLSELNKELEAERVIAITQQRWDNNIGELAGITKLGERIKSGKAYADA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 SRYAAYLANNLKFLSELNKELEAERVIAITQQRWDNNIGELAGITKLGERIKSGKAYADA 480

Qy        481 FEDGKKVEAGSNITLDAKTGIIDISNSNGKKTQALHFTSPLLTAGTESRERLTNGKYSYI 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 FEDGKKVEAGSNITLDAKTGIIDISNSNGKKTQALHFTSPLLTAGTESRERLTNGKYSYI 540

Qy        541 NKLKFGRVKNWQVTDGEASSKLDFSKVIQRVAETEGTDEIGLIVNAKAGNDDIFVGQGKM 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 NKLKFGRVKNWQVTDGEASSKLDFSKVIQRVAETEGTDEIGLIVNAKAGNDDIFVGQGKM 600

Qy        601 NIDGGDGHDRVFYSKDGGFGNITVDGTSATEAGSYTVNRKVARGDIYHEVVKRQETKVGK 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 NIDGGDGHDRVFYSKDGGFGNITVDGTSATEAGSYTVNRKVARGDIYHEVVKRQETKVGK 660

Qy        661 RTETIQYRDYELRKVGYGYQSTDNLKSVEEVIGSQFNDVFKGSKFNDIFHSGEGDDLLDG 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 RTETIQYRDYELRKVGYGYQSTDNLKSVEEVIGSQFNDVFKGSKFNDIFHSGEGDDLLDG 720

Qy        721 GAGDDRLFGGKGNDRLSGDEGDDLLDGGSGDDVLNGGAGNDVYIFRKGDGNDTLYDGTGN 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GAGDDRLFGGKGNDRLSGDEGDDLLDGGSGDDVLNGGAGNDVYIFRKGDGNDTLYDGTGN 780

Qy        781 DKLAFADANISDIMIERTKEGIIVKRNDHSGSINIPRWYITSNLQNYQSNKTDHKIEQLI 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 DKLAFADANISDIMIERTKEGIIVKRNDHSGSINIPRWYITSNLQNYQSNKTDHKIEQLI 840

Qy        841 GKDGSYITSDQIDKILQDKKDGTVITSQELKKLADENKSQKLSASDIASSLNKLVGSMAL 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GKDGSYITSDQIDKILQDKKDGTVITSQELKKLADENKSQKLSASDIASSLNKLVGSMAL 900

Qy        901 FGTANSVSSNALQPITQPTQGILAPSV 927
              |||||||||||||||||||||||||||
Db        901 FGTANSVSSNALQPITQPTQGILAPSV 927
```

RESULT 2
US-09-884-696-3
; Sequence 3, Application US/09884696
; Publication No. US20030035809A1
; GENERAL INFORMATION:
; APPLICANT: GEORGE, LISLE W
; APPLICANT: ANGELOS, JOHN A
; APPLICANT: HESS, JOHN F
; TITLE OF INVENTION: MORAXELLA BOVIS CYTOTOXIN, CYTOTOXIN GENE, ANTIBODIES
; TITLE OF INVENTION: AND VACCINES FOR PREVENTION AND TREATMENT OF MORAXELLA
; TITLE OF INVENTION: BOVIS INFECTIONS
; FILE REFERENCE: 481.06
; CURRENT APPLICATION NUMBER: US/09/884,696
; CURRENT FILING DATE: 2001-06-19
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
;   LENGTH: 953
;   TYPE: PRT
;   ORGANISM: Pasteurella haemolytica

US-09-884-696-3

```
  Query Match             50.0%;  Score 2325;  DB 11;  Length 953;
  Best Local Similarity   49.9%;  Pred. No. 2.5e-158;
  Matches  464;  Conservative  175;  Mismatches  264;  Indels   26;  Gaps   13;

Qy          8 KSNIQAGLNSTKSGLKNLYLAIPKD--YDPQKGGTLNDFIKAADELGIARLAEEPNHTET 65
Db         28 QSLTQAG-SSLKTGAKKIILYIPQNYQYDTEQGNGLQDLVKAAEELGIEVQREERNNIAT 86

Qy         66 AKKSVDTVNQFLSLTQTGIAISATKLEKFLQKHSTNKLAKGLDSVENIDRKLGKASNVLS 125
Db         87 AQTSLGTIQTAIGLTERGIVLSAPQIDKLLQK---TKAGQALGSAESIVQNANKAKTVLS 143

Qy        126 TLSSFLGTALAGIELDSLIKKGDAAPDALAKASIDLINEIIGNLSQSTQTIEAFSSQLAK 185
Db        144 GIQSILGSVLAGMDLDEAL-QNNSNQHALAKAGLELTNSLIENIANSVKTLDEFGEQISQ 202

Qy        186 LGSTISQAKGFSNIGNKLQNL-NFSKTNLGLEIITGLLSGISAGFALADKNASTGKKVAA 244
Db        203 FGSKLQNIKGLGTLGDKLKNIGGLDKAGLGLDVISGLLSGATAALVLADKNASTAKKVGA 262

Qy        245 GFELSNQVIGNVTKAISSYVLAQRVAAGLSTTGAVAALITSSIMLAISPLAFMNAADKFN 304
Db        263 GFELANQVVGNITKAVSSYILAQRVAAGLSSTGPVAALIASTVSLAISPLAFAGIADKFN 322

Qy        305 HANALDEFAKQFRKFGYDGDHLLAEYQRGVGTIEASLTTISTALGAVSAGVSAAAVGSAV 364
Db        323 HAKSLESYAERFKKLGYDGDNLLAEYQRGTGTIDASVTAINTALAAIAGGVSAAAAGSVI 382

Qy        365 GTPIALLVAGVTGLISGILEASKQAMFESVANRLQGKILEWEKQNGGQNYFDKGYDSRYA 424
Db        383 ASPIALLVSGITGVISTILQYSKQAMFEHVANKIHNKIVEWEKNNHGKNYFENGYDARYL 442

Qy        425 AYLANNLKFLSELNKELEAERVIAITQQRWDNNIGELAGITKLGERIKSGKAYADAFEDG 484
Db        443 ANLQDNMKFLLNLNKELQAERVIAITQQQWDNNIGDLAGISRLGEKVLSGKAYVDAFEEG 502

Qy        485 KKVEAGSNITLDAKTGIIDISNSNGKKTQALHFTSPLLTAGTESRERLTNGKYSYINKLK 544
Db        503 KHIKADKLVQLDSANGIIDVSNSGKAKTQHILFRTPLLTPGTEHRERVQTGKYEYITKLN 562

Qy        545 FGRVKNWQVTDGEASSKLDFSKVIQRV-------AETEGTDEIGLIVNAKAGNDDIFVGQ 597
Db        563 INRVDSWKITDGAASSTFDLTNVVQRIGIELDNAGNVTKTKETKIIAKLGEGDDNVFVGS 622

Qy        598 GKMNIDGGDGHDRVFYSKDGGFGNITVDGTSATEAGSYTVNRKVARGDIYHEVVKRQETK 657
Db        623 GTTEIDGGEGYDRVHYSR-GNYGALTIDATKETEQGSYTVNRFVETGKALHEVTSTHTAL 681

Qy        658 VGKRTETIQYRDYELRKVGYGYQSTDNLKSVEEVIGSQFNDVFKGSKFNDIFHSGEGDDL 717
Db        682 VGNREEKIEYR-HSNNQHHAGYYTKDTLKAVEEIIGTSHNDIFKGSKFNDAFNGGDGVDT 740

Qy        718 LDGGAGDDRLFGGKGNDRLSGDEGDDLLDGGSGDDVLNGGAGNDVYIFRKGDGNDTLYDG 777
Db        741 IYGNDGNDRLFGGKGDDILDGGNGDDFIDGGKGNDLLHGGKGDDIFVHRKGDGNDIITDS 800

Qy        778 TGNDKLAFADANISDIMIERTKEGIIVKRNDHSGSINIPRWY----ITSNLQNYQSNKTD 833
Db        801 DGNDKLSFSDSNLKDLTFEKVKHNLVI-TNSKKEKVTIQNWFREADFAKEVPNYKATK-D 858

Qy        834 HKIEQLIGKDGSYITSDQIDKILQDKKDGTVITSQELKKLADENKSQKLSASDIASSLNK 893
Db        859 EKIEEIIGQNGERITSKQVDDLI--AKGNGKITQDELSKVVDNYELLKHS-KNVTNSLDK 915

Qy        894 LVGSMALFGTANSVSSNALQPITQPTQGI 922
Db        916 LISSVSAFTSSNDSRNVLVAPTSMLDQSL 944
```

RESULT 3
US-09-884-696-4

```
; Sequence 4, Application US/09884696
; Publication No. US20030035809A1
; GENERAL INFORMATION:
;  APPLICANT: GEORGE, LISLE W
;  APPLICANT:  ANGELOS, JOHN A
;  APPLICANT:  HESS, JOHN F
;  TITLE OF INVENTION: MORAXELLA BOVIS CYTOTOXIN, CYTOTOXIN GENE, ANTIBODIES
;  TITLE OF INVENTION:  AND VACCINES FOR PREVENTION AND TREATMENT OF MORAXELLA
;  TITLE OF INVENTION:  BOVIS INFECTIONS
;  FILE REFERENCE: 481.06
;  CURRENT APPLICATION NUMBER: US/09/884,696
;  CURRENT FILING DATE:  2001-06-19
;  NUMBER OF SEQ ID NOS: 41
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
;   LENGTH: 956
;   TYPE: PRT
;   ORGANISM: Actinobacillus pleuropneumoniae
US-09-884-696-4

  Query Match             48.1%;  Score 2236.5;  DB 11;  Length 956;
  Best Local Similarity   47.7%;  Pred. No. 5.8e-152;
  Matches  453;  Conservative  188;  Mismatches  255;  Indels   53;  Gaps   15;

Qy          1 MSNINV--IKSNIQAGL---------------------NSTKSGLKNLYLAIPKDYDPQ 36
Db          1 MSKITLSSLKSSLQQGLKNGKNKLNQAGTTLKNGLTQTGHSLQNGAKKLILYIPQGYDSG 60

Qy         37 KGGTLNDFIKAADELGIARLAEEPNHTETAKKSVDTVNQFLSLTQTGIAISATKLEKFLQ 96
Db         61 QGNGVQDLVKAANDLGIEVWREERSNLDIAKTSFDTTQKILGFTDRGIVLFAPQLDNLLK 120

Qy         97 KHSTNKLAKGLDSVENIDRKLGKASNVLSTLSSFLGTALAGIELDSLIKKGDAAPDALAK 156
Db        121 KNP--KIGNTLGSASSISQNIGKANTVLGGIQSILGSVLSGVNLNELLQNKDPNQLELAK 178

Qy        157 ASIDLINEIIGNLSQSTQTIEAFSSQLAKLGSTISQAKGFSNIGNKLQNL-NFSKTNLGL 215
Db        179 AGLELTNELVGNIASSVQTVDAFAEQISKLGSHLQNVKGLGGLSNKLQNLPDLGKASLGL 238

Qy        216 EIITGLLSGISAGFALADKNASTGKKVAAGFELSNQVIGNVTKAISSYVLAQRVAAGLST 275
Db        239 DIISGLLSGASAGLILADKEASTEKKAAAGVEFANQIIGNVTKAVSSYILAQRVASGLSS 298

Qy        276 TGAVAALITSSIMLAISPLAFMNAADKFNHANALDEFAKQFRKFGYDGDHLLAEYQRGVG 335
Db        299 TGPVAALIASTVALAVSPLSFLNVADKFKQADLIKSYSERFQKLGYDGDRLLADFHRETG 358

Qy        336 TIEASLTTISTALGAVSAGVSAAAVGSAVGTPIALLVAGVTGLISGILEASKQAMFESVA 395
Db        359 TIDASVTTINTALAAISGGVGAASAGSLVGAPVALLVAGVTGLITTILEYSKQAMFEHVA 418

Qy        396 NRLQGKILEWEKQNGGQNYFDKGYDSRYAAYLANNLKFLSELNKELEAERVIAITQQRWD 455
Db        419 NKVHDRIVEWEKKH-NKNYFEQGYDSRHLADLQDNMKFLINLNKELQAERVVAITQQRWD 477

Qy        456 NNIGELAGITKLGERIKSGKAYADAFEDGKKVEAGSNITLDAKTGIIDISNSNGKKTQAL 515
Db        478 NQIGDLAAISRRTDKISSGKAYVDAFEEGQHQSYDSSVQLDNKNGIINISNTN-RKTQSV 536

Qy        516 HFTSPLLTAGTESRERLTNGKYSYINKLKFGRVKNWQVTDGEASSKLDFSKVIQRVA--- 572
Db        537 LFRTPLLTPGEENRERIQEGKNSYITKLHIQRVDSWTVTDGDASSSVDFTNVVQRIAVKF 596

Qy        573 -------ETEGTDEIGLIVNAKAGNDDIFVGQGKMNIDGGDGHDRVFYSKDGGFGNITVD 625
Db        597 DDAGNIIESKDTK---IIANLGAGNDNVFVGSSTTVIDGGDGHDRVHYSR-GEYGALVID 652

Qy        626 GTSATEAGSYTVNRKVARGDIYHEVVKRQETKVGKRTETIQYRDYELRKVGYGYQSTDNL 685
Db        653 ATAETEKGSYSVKRYVGDSKALHETIATHQTNVGNREEKIEYR-REDDRFHTGYTVTDSL 711

Qy        686 KSVEEVIGSQFNDVFKGSKFNDIFHSGEGDDLLDGGAGDDRLFGGKGNDRLSGDEGDDLL 745
Db        712 KSVEEIIGSQFNDIFKGSQFDDVFHGGNGVDTIDGNDGDDHLFGGAGDDVIDGGNGNFL 771

Qy        746 DGGSGDDVLNGGAGNDVYIFRKGDGNDTLYDGTGNDKLAFADANISDIMIERTKEGIIVK 805
Db        772 VGGTGNDIISGGKDNDIYVHKTGDGNDSITDSGGQDKLAFSDVNLKDLTFKKVDSSLEI- 830

Qy        806 RNDHSGSINIPRWY----ITSNLQNYQSNKTDHKIEQLIGKDGSYITSDQIDKILQDKKD 861
Db        831 INQKGEKVRIGNWFLEDDLASTVANYKAT-NDRKIEEIIGKGGERITSEQVDKLI--KEG 887

Qy        862 GTVITSQELKKLADENKSQKLSASDIASSLNKLVGSMALFGTANSVSSN 910
Db        888 NNQISAEALSKVVNDYNTSK-DRQNVSNSLAKLISSVGSFTSSSDFRNN 935


RESULT 4
US-09-884-696-5
; Sequence 5, Application US/09884696
; Publication No. US20030035809A1
; GENERAL INFORMATION:
;  APPLICANT: GEORGE, LISLE W
;  APPLICANT:  ANGELOS, JOHN A
;  APPLICANT:  HESS, JOHN F
;  TITLE OF INVENTION: MORAXELLA BOVIS CYTOTOXIN, CYTOTOXIN GENE, ANTIBODIES
;  TITLE OF INVENTION:  AND VACCINES FOR PREVENTION AND TREATMENT OF MORAXELLA
;  TITLE OF INVENTION:  BOVIS INFECTIONS
;  FILE REFERENCE: 481.06
;  CURRENT APPLICATION NUMBER: US/09/884,696
;  CURRENT FILING DATE:  2001-06-19
;  NUMBER OF SEQ ID NOS: 41
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
;   LENGTH: 1023
;   TYPE: PRT
;   ORGANISM: Escherichia coli
US-09-884-696-5

  Query Match             39.7%;  Score 1845;  DB 11;  Length 1023;
  Best Local Similarity   41.8%;  Pred. No. 9.2e-124;
  Matches  421;  Conservative  156;  Mismatches  311;  Indels  118;  Gaps   23;

Qy          8 KSNIQAGLNSTKSGLKNLYLAIPKDYDPQKGGTLNDFIKAADELGIARLAEEPNHTETAK 67
Db         32 KDALKKAAEQTRNAGNRLILLIPKDYKGQ-GSSLNDLVRTADELGIEVQYDEKNGTAITK 90

Qy         68 KSVDTVNQFLSLTQTGIAISATKLEKFLQKH--STNKLAKGLDSVENIDRKLGKASNVLS 125
Db         91 QVFGTAEKLIGLTERGVTIFAPQLDKLLQKYQKAGNKLG---GSAENIGDNLGKAGSVLS 147

Qy        126 TLSSFLGTALAGIELDSLIKK----GDAAPDALAKASIDLINEIIGNLSQSTQTIEAFSS 181
Db        148 TFQNFLGTALSSMKIDELIKKQKSGGNVSSSELAKASIELINQLV-DTAASLNNVNSFSQ 206

Qy        182 QLAKLGSTISQAKGFSNIGNKLQNL-NFSKTNLGLEIITGLLSGISAGFALADKNASTGK 240
Db        207 QLNKLGSVLSNTKHLNGVGNKLQNLPNLDNIGAGLDTVSGILSAISASFILSNADADTGT 266

Qy        241 KVAAGFELSNQVIGNVTKAISSYVLAQRVAAGLSTTGAVAALITSSIMLAISPLAFMNAA 300
Db        267 KAAAGVELTTKVLGNVGKGISQYIIAQRAAQGLSTSAAAAGLIASVVTLAISPLSFLSIA 326

Qy        301 DKFNHANALDEFAKQFRKFGYDGDHLLAEYQRGVGTIEASLTTISTALGAVSAGVSAAAV 360
Db        327 DKFKRANKIEEYSQRFKKLGYDGDSLLAAFHKETGAIDASLTRISTVLASVSSGISAAAT 386

Qy        361 GSAVGTPIALLVAGVTGLISGILEASKQAMFESVANRLQGKILEWEKQNGGQNYFDKGYD 420
Db        387 TSLVGAPVSALVGAVTGIISGILEASKQAMFEHVASKMADVIAEWEKKH-GKNYFENGYD 445

Qy        421 SRYAAYLANNLKFLSELNKELEAERVIAITQQRWDNNIGELAGITKLGERIKSGKAYADA 480
Db        446 ARHAAFLEDNFKILSQYNKEYSVERSVLITQQHWDTLIGELAGVTRNGDKTLSGKSYIDY 505
```

```
Qy        481 FEDGKKVEAG----SNITLDAKTGIIDISNSNGKKTQALHFTSPLLTAGTESRERLTNGK 536
Db        506 YEEGKRLEKKPDEFQKQVFDPLKGNIDLSDS--KSSTLLKFVTPLLTPGEEIRERRQSGK 563

Qy        537 YSYINKLKFGRVKNWQVTD-GEASSKLDFSKVIQRVAETEGT-DEIGLIVNAKAGNDDIF 594
Db        564 YEYITELLVKGVDKWTVKGVQDKGSVYDYSNLIQHASVGNNQYREIRIESHLGDGDDKVF 623

Qy        595 VGQGKMNIDGGDGHDRVFYSK-DGGFGNITVDGTSATEAGSYTVNRKVARGD--IYHEVV 651
Db        624 LSAGSANIYAGKGHDVVYYDKTDTGY--LTIDGTKATEAGNYTVTR-VLGGDVKVLQEVV 680

Qy        652 KRQETKVGKRTETIQYRDYELRKV-GYGYQSTDNLKSVEEVIGSQFNDVFKGSKFNDIFH 710
Db        681 KEQEVSVGKRTEKTQYRSYEFTHINGKNLTETDNLYSVEELIGTTRADKFFGSKFADIFH 740

Qy        711 SGEGDDLLDGGAGDDRLF---------------------------------------- 728
Db        741 GADGDDHIEGNDGNDRLYGDKGNDTLSGGNGDDQLYGGDGNDKLIGGAGNNYLNGGDGDD 800

Qy        729 --------------GGKGNDRLSGDEGDDLLDGGSGDDVLNGGAGNDVYIFRKGDGNDT 773
Db        801 ELQVQGNSLAKNVLSGGKGNDKLYGSEGADLLDGGEGNDLLKGGYGNDIYRYLSGYGHHI 860

Qy        774 LYDGTG-NDKLAFADANISDIMIERTKEGIIVKRND-------HSGSINIPRWYITSNLQ 825
Db        861 IDDDGGKDDKLSLADIDFRDVAFRREGNDLIMYKAEGNVLSIGHKNGITFKNWF-----E 915

Qy        826 NYQSNKTDHKIEQLIGKDGSYITSDQIDKILQDKK-----------DGTVITSQ------ 868
Db        916 KESGDISNHQIEQIFDKDGRVITPDSLKKALEYQQSNNKASYVYGNDALAYGSQGNLNPL 975

Qy        869 --ELKKLADENKSQKLSASDIASSLNKLVGSMALFGTA-NSVSSNA 911
Db        976 INEISKIISAAGNFDVKEERAAASLLQLSGNASDFSYGRNSITLTA 1021
```

RESULT 5
US-09-305-924-13
; Sequence 13, Application US/09305924A
; Publication No. US20030091579A1
; GENERAL INFORMATION:
; APPLICANT: Jack G. Manns
; APPLICANT: Stephen D. Acres
; APPLICANT: Richard Harland
; TITLE OF INVENTION: METHODS OF RAISING ANIMALS FOR MEAT PRODUCTION
; FILE REFERENCE: 9001-0048
; CURRENT APPLICATION NUMBER: US/09/305,924A
; CURRENT FILING DATE: 1999-05-05
; EARLIER APPLICATION NUMBER: US 60/084,217
; EARLIER FILING DATE: 1998-05-05
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 13
; LENGTH: 695
; TYPE: PRT
; ORGANISM: GnRH
US-09-305-924-13

Query Match 19.4%; Score 901; DB 11; Length 695;
Best Local Similarity 38.8%; Pred. No. 3.9e-56;
Matches 221; Conservative 99; Mismatches 165; Indels 84; Gaps 14;

```
Qy         14 GLNSTKSGLKNLYLAIPKD--YDPQKGGTLNDFIKAADELGIARLAEEPNHTETAKKSVD 71
Db        109 GSSFPKTGAKKIILYIPQNYQYDTEQGNGLQDLVKAAEELGIEVQREERNNIATAQTSLG 168

Qy         72 TVNQFLSLTQTGIAISATKLEKFLQKHSTNKLAKGLDSVENIDRKLGKASNVLSTLSSFL 131
Db        169 TIQTAIGLTERGIVLSAPQIDKLLQK---TKAGQALGSAESIVQNANKAKTVLSGIQSIL 225

Qy        132 GTALAGIELDSLIKKGDAAPDALAKASIDLINEIIGNLSQSTQTIEAFSSQLAKLGSTIS 191
Db        226 GSVLAGMDLDEAL-QNNSNQHALAKAGLELTNSLIENIANSVKTLDEFGEQISQFGSKLQ 284

Qy        192 QAKGFSNIGNKLQNL-NFSKTNLGLEIITGLLSGISAGFALADKNASTGKKVAAGFELSN 250
Db        285 NIKGLGTLGDKLKNIGGLDKAGLGLDVISGLLSGATAALVLADKNASTAKKVGAGFELAN 344

Qy        251 QVIGNVTKAISSYVLAQRVAAGLSTTGAVAALITSSIMLAISPLAFMNAADKFNHANALD 310
Db        345 QVVGNITKAVSSYILAQRVAAGLSSTGPVAALIASTVSLAISPLAFAGIADKFNHAKSLE 404

Qy        311 EFAKQFRKFGYDGDHLLAEYQRGVGTIEASLTTISTALGAVSAGVSAAAVGSAVGTPIAL 370
Db        405 SYAERFKKLGYDGDNLLAEYQRGTGTIDASVTAINTALAAIAGGVSAAA----------- 453

Qy        371 LVAGVTGLISGILEASKQAMFESVANRLQGKILEWEKQNGGQNYFDKGYDSRYAAYLANN 430
Db        454 --ADLT-------------FEKVKHNLVITNSKKEKVT-IQNWF-------------- 481

Qy        431 LKFLSELNKELEAERVIAITQQRWDNNIGELAGITKLGERIKSGKAYADAFEDGKKVEAG 490
Db        482 --------READFAKEVPNYKATKDEKIEEIIG--QNGERITSKQV------DDLIAKGN 525

Qy        491 SNITLDAKTGIID----------ISNSNGKKTQALHFTSPLLTAGTESRERL---TNGKY 537
Db        526 GKITQDELSKVVDNYELLKHSKNVTNSLDKLISSV----SAFTSSNDSRNVLVAPTSMLD 581

Qy        538 SYINKLKFGR-VKNWQVTDGEASSKLDFS 565
Db        582 QSLSSLQFARGSQHWSYGLRPGSGSQDWS 610
```

RESULT 6
US-10-193-950A-2
; Sequence 2, Application US/10193950A
; Publication No. US20030104002A1
; GENERAL INFORMATION:
; APPLICANT: Sparling , P. Frederick
; APPLICANT: Thompson, Stuart
; TITLE OF INVENTION: Antigenic iron repressible proteins from N. meningitidis related
; TITLE OF INVENTION: hemolysin family of toxins
; FILE REFERENCE: 381-39 CIP/CONII/CPA/CON
; CURRENT APPLICATION NUMBER: US/10/193,950A
; CURRENT FILING DATE: 2002-12-23
; PRIOR APPLICATION NUMBER: 09/045,177
; PRIOR FILING DATE: 1998-03-20
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 1115
; TYPE: PRT
; ORGANISM: Neisseria Meningitidis
US-10-193-950A-2

Query Match 8.8%; Score 408.5; DB 15; Length 1115;
Best Local Similarity 20.4%; Pred. No. 2.1e-20;
Matches 239; Conservative 171; Mismatches 364; Indels 399; Gaps 51;

```
Qy         40 TLNDFIKAADELGIARLAEEPNHTETAKKSVDTVNQFLSL-------TQTG--------- 83
Db         55 TVNNAIQAADSF-LSSIGRDNKITNTASLLASLDNIFLNLRNVSRDIRETGKFKPNDIQQ 113

Qy         84 ----IAISA--------TKLEKFLQ-KHSTNKLAKGLDSVENIDRKLGKASNVLSTLSSF 130
Db        114 AIGDIFIAAGDGLQYIKQQTEAMAQSKFLPTKLKTGLNDV--LNSRMLKSSTVLQHELNY 171

Qy        131 LGTALA--GIEL--DSLIKKGDAAPDALAK--ASIDLINEIIGNLSQSTQTIEAFSSQLA 184
Db        172 LGFKIKDYGNERLGESIMNIDDFTPSKIANFFADPDTYSNVLEEVSRFIYSLVPDDANPW 231

Qy        185 K-----LGSTISQ------------------------------------------- 192
Db        232 KGGEDYIGRGISEWGELLEKWYKQDFLPYLEKEWDQFPKFEDWLPEFPEWAREWLKLDPK 291
```

```
Qy    193 -----------------------AKGFSNIGNKLQNLNFSKTNLGLEIITGLLSGISAG 228
Db    292 RSGKYHVYDPLALDLDGDGIETVAAKGFA-------GALFDHRNQGIRTATGWVSA-DDG 343

Qy    229 FALADKNASTGKKVAAGFELSNQVIGNVTKAISSYVLAQRVAAGLSTTGAVAALITSSIM 288
Db    344 LLVRDLN------------GNGIIDNGAELFGDNT---KLADGSFAKHGYAALAELD-- 385

Qy    289 LAISPLAFMNAADKFNHANALDEFAKQFRKFGYDG----DHLLAEYQRGVGTIEASLTTI 344
Db    386 --------SNGDNIINAADAAFQTLRVWQDLNQDGISQANELRTLEELGIQSLDLAYKDV 437

Qy    345 STALGAVSAGVSAAAVGSAVGTPIALLVAGVTGLISGILEASKQAMFESVANRLQGKILE 404
Db    438 NKNLG---NGNTLAQQGSYTKTD------GTTAKMGDLLLAA-----DNLHSRFKDKVEL 483

Qy    405 WEKQNGGQNYFDKG--YDSRYAAYLANNLKFLSELNKELEA-ERVIAITQ---QRW---D 455
Db    484 TAEQAKAANLAGIGRLRDLREAAALSGDLANMLKAYSAAETKEAQLALLDNLIHKWAETD 543

Qy    456 NNIGELA-------------GITKLGERI-----------KSGKAYADAFEDGKKVEA 489
Db    544 SNWGKKSPMRLSTDWTQTANEGIALTPSQVAQLKKNALVSLSDKAKAAIDAARDRIAV-- 601

Qy    490 GSNITLDAKTG---------------IIDISNSN------------------------ 508
Db    602 -----LDAYTGQDSSTLYYMSEEDALNIVKVTNDTYDHLAKNIYQNLLFQTRLQPYLNQI 656

Qy    509 ----GKKTQALHFTSPLLTAGTESRERLTNGKYSYIN---KLKFGRVKNW---------Q 552
Db    657 SFKMENDTFTLDF-SGLVQAFNHVKE--TNPQKAFVDLAEMLAYGELRSWYEGRRLMADY 713

Qy    553 VTDGEASSKL-DFSKVIQR-----VAETEGTDEIGLIVN-----------AKAGNDDIF 594
Db    714 VEEAKKAGKFEDYQKVLGQETVALLAKTSGTQADDILQNVGFGHNKNVSLYGNDGNDTLI 773

Qy    595 VGQGKMNIDGGDGHDRVFYSKDGGFGNITV---------------DGTSA-----TEAG 633
Db    774 GGAGNDYLEGGSGSDTYVFGK--GFGQDTVYNYDYATGRKDIIRFTDGITADMLTFTREG 831

Qy    634 SY-TVNRKVARGDIYHEVVKRQETKVGKRTETIQYRDYELRKVG--------------- 676
Db    832 NHLLIKAKDDSGQVTVQSYFQNDGSGAYRIDEIHFDNGKVLDVATVKELVQQSTDGSDRL 891

Qy    677 YGYQSTDNLK---SVEEVIGSQFNDVFKGSKFNDIFHSGEGDDLLDGGAGDDRLF----- 728
Db    892 YAYQSGSTLNGGLGDDYLYGADGNDLLNGDAGNDSIYSGNGNDTLDGGEGNDALYGYNGN 951

Qy    729 ----GGKGNDRLSGDEGDDLLDGGSGDDVLNGGAGNDVYIFRKGDGNDTLYD---GTGND 781
Db    952 DALNGGEGNDHLNGEDGNDTLIGGAGNDYLEGGSGSDTYVFGEGFGQDTVYNYHVDKNSD 1011

Qy    782 KLAFADANISDIMIERTKEGIIVKRNDHSGSINIPRWYITSNLQNYQSNKTDHKIEQLIG 841
Db   1012 TMHFKGFKAADVHFIRSGSDLVLSASEQD-NVRISGFFYGEN---------HRVDTFVF 1060

Qy    842 KDGSYITSDQIDKILQDKKDGTVITSQELKKLADENKSQKLSASDIASSL---NKLVGSM 898
Db   1061 DDAA---------------------------------ISNPDFAKYINAGNNLVQSM 1084

Qy    899 ALFG-----TANSVSSNALQPITQPTQGILAPS 926
Db   1085 SVFGSNTAATGGNVDAN-IQSVQQPL--LVTPS 1114
```

RESULT 7
US-10-369-493-20166
; Sequence 20166, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 20166
; LENGTH: 1017
; TYPE: PRT
; ORGANISM: No. US20030233675A1toc punctiforme
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(1017)
; OTHER INFORMATION: unsure at all Xaa locations
US-10-369-493-20166

```
  Query Match             6.4%;  Score 295.5;  DB 12;  Length 1017;
  Best Local Similarity   22.6%;  Pred. No. 2.5e-12;
  Matches  182;  Conservative   92;  Mismatches  282;  Indels  249;  Gaps   35;

Qy    120 ASNVLSTLSSFLGTALAGIEL---DSLIKKGDAAPDALAKASIDLINEIIGNLSQSTQTI 176
Db    331 ASSSNNTLDGGAGNDTLNAESPSGDNLLSGGD-GNDYLSTSGTKSDNGQLNFTSEGNNTL 389

Qy    177 EAFSSQLAKLGSTISQAKGFSNIGNKLQNLNFSKTNLGLEII-----TGLLSGISAGFAL 231
Db    390 DGGAGD-DRLYADISTGNNLLSGGDGNDTLSTSGTYTRNSFIVPSSGNNTLNG-GAGDDY 447

Qy    232 ADKNASTGKKVAAG------FELSNQVIGNVTKAISSYVLAQRVAAGLSTTGAVAALITS 285
Db    448 LYASASTGSNFFSGGDGNDSFSLS--TISSDGSPSNSYLVTQTVDGG-----------KG 494

Qy    286 SIMLAISPLAFMNAADKFNHANALDEFAKQFRKFGYDGDHLLAEYQRGVGTIEASLTTI- 344
Db    495 DDLLSFNLLPF----------------------HDSDSNVA------GGI---ITTFN 521

Qy    345 -STALGAVSAGVSAAAVG-----SAVGTPIALLVAGVTG---LISGILEASKQAMFESVA 395
Db    522 PTTNKGTITAGTYGVNYKNIEKLNISGTADDDLILGSNGNDTLSTG-------------- 567

Qy    396 NRLQGKILEWEKQNGGQNYFD--KGYDSRYAAYLANNLKFLSELNKELEAERVIAITQQR 453
Db    568 ------------NGGKDTIDGGKGDDVISIDYSNATSKIITTFNATTNTGLITAGTYQI 614

Qy    454 WDNNIGELAGITKLGERIK-SGKAYADAF-----EDGKKVEAGSNITLDAKTG--IIDI- 504
Db    615 SYKNI----------ERLNITGTAYDDNILGNNGNDTLSTGSGGNDTINGGQGDDILSVD 664

Qy    505 --SNSNGKKTQALHFTSP-LLTAGTESRERLTNGKYSYINKLKF-GRVKNWQVTDGEASS 560
Db    665 YYTNTTGITTTYNTTTNAGVITAGS------NRVSYKNIERLDISGTVSDDYIIGNNGND 718

Qy    561 KLDFSKVIQRVAETEGTDEIGLIVNAKAGNDDIFVGQGKMNIDGGDGHDRVFY-----SK 615
Db    719 TL-----------SGGYDGNDTIIGG-AGNDVVIGGKGNDILTGGVGNDKFVYDLLESNY 766

Qy    616 DGG------FGNI----------------------------------------------- 622
Db    767 DTGTDIITDFGGIGKDSNPSQAVIASLDTLEFIDNSGFRLFSAQNLQLTQNGNNLEITFE 826

Qy    623 -----------------TVDGTSATEAGSYTVNRKVARGDIYHEVVKRQETKVGKRTET 664
Db    827 DLFNSSNSKVILQNFQLENLDGENILFPGQTTITNSI---DVF-DANSTQTSLLNKNTVT 882

Qy    665 IQYRDYELRKVGYGYQSTDNLKSVEEVIGSQFNDVFKGSKFNDIFHSGEGDDLLDGGAGD 724
Db    883 F------LNDLNNNITGFDN----------SNDVVNGQGGDDIINGLSGNDLLRGNAGN 925
```

```
Qy        725 DRLFGGKGNDRLSGDEGDDLLDGGSGDDVLNGGAGNDVYIFRKGDGNDTLYDGTGNDK-- 782
              | | || ||| | |   |:||| ||:| |:| |   |||:: |   |:| ||: | | |:
Db        926 DTLIGGAGNDTLYGGAGNDLLYGGAGSDILTGDNGNDIFAFAAGEGTDTITDFTKNNDLI 985

Qy        783 -----LAFADANIS--DIMIERTKE 800
                   |:|   : |  :|::  | |
Db        986 GLYGGLSFGQLSFSGNNILVTSTNE 1010

RESULT 8
US-10-369-493-10178
; Sequence 10178, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
;  APPLICANT: Cao, Yongwei
;  APPLICANT:  Hinkle, Gregory J.
;  APPLICANT:  Slater, Steven C.
;  APPLICANT:  Goldman, Barry S.
;  APPLICANT:  Chen, Xianfeng
;  TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
;  TITLE OF INVENTION:  PLANTS WITH IMPROVED PROPERTIES
;  FILE REFERENCE: 38-10(52052)B
;  CURRENT APPLICATION NUMBER: US/10/369,493
;  CURRENT FILING DATE:  2003-02-28
;  PRIOR APPLICATION NUMBER: US 60/360,039
;  PRIOR FILING DATE: 2002-02-21
;  NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 10178
;   LENGTH: 4327
;   TYPE: PRT
;   ORGANISM: magnetite-containing magnetic coccus
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: (1)..(4327)
;   OTHER INFORMATION: unsure at all Xaa locations
US-10-369-493-10178

  Query Match             5.7%;  Score 263.5;  DB 12;  Length 4327;
  Best Local Similarity   19.8%;  Pred. No. 4.5e-09;
  Matches  203;  Conservative  123;  Mismatches  358;  Indels  339;  Gaps   44;

Qy          1 MSNINVIKSNIQAGLN----STKSGLKNLYLAIPKDYDPQKGGTLNDFIKAADELGIARL 56
               :||| :    |   |        :    |: ||      |        || |   :     |    :|:|
Db       3165 VSNITIDSGNYVANAGGATITVNDGIANLI----TDAGTVVTGTRN--VTVTDAASMAQL 3218

Qy         57 AEEPNHTETAKKSV---DTVNQFLSLTQTGI----AISATKLEKFLQKH---------- 98
              ::  |:|  | | |   | |   :: | : :    |:|  |         |
Db       3219 SQIDNYTTGALKYVTIKDAVAALVANTNSYVTGSYAVSVTDAASMAQLSAIDQDTTGTLT 3278

Qy         99 ------------STNKLAKG-----------LDSVENID---------RKLGKASNVLS 125
                          :||                :  : :||          ::| |:  |:
Db       3279 YTKLTDAVANLVTNTNSYVTGSVNVTVSDIATISQLSSIDANTTGSVTYTQIGDAAATLA 3338

Qy        126 T---------------------LSSFLGTALAGIELDSLIKKGDAAPDALAKASIDLIN 163
              |                      |::  |      |  ||:           |:  |      ::   :
Db       3339 TNAGNYVKATIHVTVTDAATIAQLTTIDGDNTTG----SLVYTAGGVKDSAANLVVNTNS 3394

Qy        164 EIIGNLSQS-TQTIEAFSSQLAKLG----STISQAKGF-SNIGNKLQNLN---------- 207
               : | ::  | | |:     :||: :          |::    |      ::  | | | |
Db       3395 YVTGAVNVSVTDTVSI--AQLSAVDEYTTGTLTYGAGVKDSVANLLVNTNSYVTGSYAVS 3452

Qy        208 ----FSKTNLGL--EIITGLLSGISAGFALADKNASTGKKVAAGFELS---NQVIGNV-- 256
                  | ||    :  || |:        ::    |:|    |          ::      |   : |:
Db       3453 ITDVASMANLSAIDQFTTGTLNYTKLSDTVSALVANTNSYVTGSVNVTITDNASMANMSA 3512

Qy        257 ----TKAISSYVLAQRVAAGLS------TTGAVAALITSSIMLAISPLAFMNAA--DKFN 304
                  |      :|            || |:           ||:|    :|  :       |  ::||
Db       3513 IDQNTTGTLTYTKLSDTAAALAANTNSYVTGSVNVTVTDN--ATVAQLTTVDAATTGTIK 3570

Qy        305 HANALDEFAKQFRKFGY-DG---------DHLLAEYQRGVGTIEASLTTI---------- 344
               |: :|  :        | | ||             |:::|        |     :|:   |:
Db       3571 VASVVDSGSNISSNFAYVDGLGVSYVNANDNVMA-----VTVAQATDATVTIADDDVLTI 3625

Qy        345 ---------STALGAVSAGVSA-----------AAVGSAVGTPIALLVAGV--------- 375
                        ||     |:||   |                 |   |   |      | |
Db       3626 TDTSSNIQGSTFSALVTAGADATYLDATDDILTVTAAQAATTNINFTAADVVTVSDTGSN 3685

Qy        376 ------TGLISGILEASKQAMFESVANRLQGKILEWEKQNGGQNYFDKGYDSRYAAYLAN 429
                    | |    : :  |    :    :    :  : || :::          |     ||| |:
Db       3686 IQTNITTILAKDVDKLSASSALDLTDSDVTGKTVDF--SGVGDTTITMAYDSSNVAF--- 3740

Qy        430 NLKFLSELNKELEAERVIAITQQRWDNNIGELAGITKLGERIKSGKAYADAFEDGKKVEA 489
                      |          | ::::|          | || :| |                   :  :: |
Db       3741 -----GSLTSTGGGELILSVT------GTGTLATVTGL--------------SNFNQLSA 3775

Qy        490 GSNITLDA---------KTGIIDISNSNGKKTQALHFTSPLLTAGTESRERLTNGKYSYI 540
              |::|:|:           :|  ::: |        | | : :         |||     :  : | |:
Db       3776 ASDLTIDSAQVAGQTLDMSGAGNVTVSLASATTASYSSLTSTGAGTLGLQIGSTGTYT-- 3833

Qy        541 NKLKFGRVKNWQVTDGEASSKLDFSKVIQRVAETEGTDEIGLIVNAKAGNDDIFVGQGKM 600
                    |  |:|  |     |   |::                  | :  :|  |      | | | :
Db       3834 ------SVTGWEVFDSITS---DYA--------------ITVAASAVTGKTVAFSGTGAV 3870

Qy        601 NIDGGDG---------HD----------------RVFYSKDGGFGNITVDGTSATEAGS 634
                 |||              ||                    || |       |  ::|||        |
Db       3871 TIDGASSSNATFLNVTHDSGTSGDLILAVSSSTGSAVFSSGAANFDKVSVDGTLYTGITS 3930

Qy        635 YTVNRKVARGDIYHEVVKRQETKVGKRTETIQYRD----YELRKVGYGYQSTDNLKSVEE 690
                     | |     :::        ||        |:        :|   | ||  |
Db       3931 -------ANG----QIIVNGSTSVAAGISGTAYTDFIVNYDTDSLGTTYASTQGLL---- 3975

Qy        691 VIGSQFNDVFKGSKFNDIFHSGEGDDLLDGGAGDDRLFGGKGNDRLSGDEGDDLLDGGSG 750
                                            |||||||                :             || |  :|
Db       3976 ----------------------------GGAGDDRL--------IDTTSSGVLLAGEAG 3998

Qy        751 DDVLNGGAGNDVYIFRKGD--GNDTLYDGTGNDKLAFADANISDIMIERTKEGIIVKRND 808
              || |  ||||||:: :  |     | ||: |     | |  |  |      |          :: |  : :      |
Db       3999 DDTLTGGAGNDIF-YHDGSSHGLDTITDFASGDLLQIASGNAGQWTLDSTGQALTKATNS 4057

Qy        809 HSG 811
                |
Db       4058 SLG 4060

RESULT 9
US-10-238-075-749
; Sequence 749, Application US/10238075
; Publication No. US20030148324A1
; GENERAL INFORMATION:
;  APPLICANT: I.N.S.E.R.M.
;  TITLE OF INVENTION: Polynucleotides which are of nature B2/D+ A- and which are isolat
;  TITLE OF INVENTION:  E.coli, and biological uses of these polynucleotides and of thei
;  FILE REFERENCE: BLANDINE
;  CURRENT APPLICATION NUMBER: US/10/238,075
;  CURRENT FILING DATE:  2002-09-10
;  PRIOR APPLICATION NUMBER: 0003145
;  PRIOR FILING DATE: 2000-03-10
;  NUMBER OF SEQ ID NOS: 1576
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 749
;   LENGTH: 1778
;   TYPE: PRT
;   ORGANISM: Escherichia coli
US-10-238-075-749

  Query Match             5.0%;  Score 232;  DB 12;  Length 1778;
  Best Local Similarity   20.7%;  Pred. No. 2.1e-07;
  Matches  214;  Conservative  134;  Mismatches  417;  Indels  268;  Gaps   45;

Qy          6 VIKSNIQAGLNSTKSGLKNLYLAIPKDYDP---------QKGGTLNDFIK---AADELG 52
               :| : :  ||  |:        :   |  ||                || | |:     |:  ||
Db         36 LISALVAGGLLSSFGASADNYTGQPTDYGDGSAGDGWVAIGKGAKANTFMNTSGASTALG 95

Qy         53 IARLAEEPNHTETAKKSVDTVNQFLSLTQTGIAISATKLEKFLQKHSTNKLAKGLDSVEN 112
                 :||      :    |:: |    ::    |:: |                    :| |  || |
Db         96 YDAIAEGEYSSAIGSKTLATGGASMAF---GVSAKAM---------GDRSVALGASSVAN 143
```

```
Qy        113 IDRKLGKASNVLSTLSSFLGTALAGIELDSLIKKGDAAPDALAKASIDLINEIIGNLSQS 172
Db        144 GDRSM--AFGRYAKTNGFTSLAIG----DSSLADGEK------------TIALGNTAKA 184

Qy        173 TQTIEAFSSQLAKLGSTISQAKGFSNIGNKLQNLNFSKTNLGLEIITGLLSGISAGFALA 232
Db        185 YEIMSIALGDNANASKEYAMALGASSKAGGADSLAFGRKSTANS--TGSLA-IGADSSSS 241

Qy        233 DKNA-STGKKVAAGFELSNQVIGNVTKAISSYVLAQRVAAGLSTTGAVAALITSSIMLAI 291
Db        242 NDNAIAIGNKTQA-LGVNSMALGNASQASGESSIALGNTSEASEQNAI-ALGQGSIASKV 299

Qy        292 SPLAF-MNAADKFNHANALDE----------FAKQFRKFGYDGDHLLAEYQRGVGTIEA 339
Db        300 NSIALGSNSLSSGENAIALGEGSAAGGSNSLAFGSQSRANGNDSVAI------GVGAAAA 353

Qy        340 SLTTISTALGAVSAGVSAAAVG-SAVGTPIALLVAGV-----TGLISGILEASKQAMFES 393
Db        354 TDNSVAIGAGSTTDASNTVSVGNSATKRKIVNMAAGAISNTSTDAING---SQLYTISDS 410

Qy        394 VANRLQGKILEWEKQNGGQNYFDKG------YDSRYAAYLANNLKFLSELNKELEAERVI 447
Db        411 VAKRL---------GGGATVGSDGTVTAVSYALRSGTY--------------------- 439

Qy        448 AITQQRWDNNIGE-LAGITKLGERIKSGKAYADAFEDGKKVEAGSNITLDAK-------T 499
Db        440 --------NNVGDALSGID--NNTLQWNKT-AGAFSANHGANATNKITNVAKGTVSATST 488

Qy        500 GIIDISNSNGKKTQALHFTSPLLTA--GTESRERLTN--------GKYSYINKLKFGRVK 549
Db        489 DVVNGSQLYDLQQDALLWNGTAFSAAHGTEATSKITNVTAGNLTAGSTDAVNGSQL-KTT 547

Qy        550 NWQVT-------------------DGEASSKLDFSKVIQRVAETEGTDEIGLIVNAKAG 589
Db        548 NDNVTTNTTNIATNTTNITNLTDAVNGLGDDSLLWNKAAGAFSAAHGTEATSKITNVTAG 607

Qy        590 N------DDIFVGQGKMNID---------------------GGDGHDRVFYSKDGGFG 620
Db        608 NLTAGSTDAVNGSQLKTTNDNVTTNTTNIATNTTNITNLTDAVNGLGDDSLLWNKTAGAF 667

Qy        621 NITVDGTSATE------AGSYTVNRKVARGDIYHEVVKRQETKVGKRTETIQYRDYELRK 674
Db        668 S-AAHGTDATSKITNVTAGNLTAGSTDA---VNGSQLKTTNDNVTTNTTNIA-------- 715

Qy        675 VGYGYQSTDNLKSVEEVIGSQFNDVFKGSKFNDIFHSGEGDD----LLDGGAGDDRLFGG 730
Db        716 -----TNTTNITNLTDAVNGLGDDSLLWNKTAGAFSAAHGTDATSKITNVKAGD--LTAG 768

Qy        731 KGNDRLSGDEGDDLLDGGSGD--------DVLNGGAGNDVYIFRKGDGNDTLYDGTG--- 779
Db        769 -STDAVNGSQLKTTNDNVSTNTTNITNLTDAVN-GLGDDSLLWNKTAGAFSAAHGTDATS 826

Qy        780 ---NDKLAFADANISDIMIERTKEGIIVKRNDHSGSINIPRWYITSNLQNYQSNKTDHKI 836
Db        827 KITNVKAGDLTAGSTDAV-----NGSQLKTTNDNVSTN------TTNITNLTDSVGDLKD 875

Qy        837 EQLIGKDGSYITSDQIDKILQDKKDGTVITSQELKKLADENKSQKLSASDIASSLNKLVG 896
Db        876 DSLLWNKAAGAFS---------AAHGTEATS----------KITNLLAGKISSNSTDAIN 916

Qy        897 SMALFGTANSVSS 909
Db        917 GSQLYGVADSFTS 929
```

RESULT 10
US-10-369-493-20175
; Sequence 20175, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 20175
; LENGTH: 499
; TYPE: PRT
; ORGANISM: No. US20030233675A1toc punctiforme
US-10-369-493-20175

Query Match 5.0%; Score 230.5; DB 12; Length 499;
Best Local Similarity 23.0%; Pred. No. 3.9e-08;
Matches 122; Conservative 59; Mismatches 169; Indels 181; Gaps 24;

```
Qy        409 NG--GQN--YFDKGYDSRYAAYLANNLKFLSELNKELEAERVIAITQQRWDNNIGELAGI 464
Db         58 NGEEGNNSLYGEAGNDSFYAANSSGN-----------------------NLLDGGAGN 92

Qy        465 TKLGERIKSGKAYADAFEDGKKVEAGSNITLDAKTG--IIDISNSNGKKTQALHFTSPLL 522
Db         93 DELYVYSSTG----------------NNTLKGGSGNDTLDISFSYGKN---------IL 126

Qy        523 TAGTESRERLTNGKYSY-INKLKFGRVKNWQVTD----------GEASSKLDFS---KVI 568
Db        127 DGGSED-DRL-SARYAVGNNTLRGGTGNDYIAADLSAGNNVLDGGDGNDTLDINYQWKGK 184

Qy        569 QRVAETEGTDEI-------GLIVNAKAGNDDIFVGQGK--------MNIDGGDGHDRVFY 613
Db        185 NTVSGGNGADTFYAYGVQGANTLNGDSGDDSFYISSPETVLYSLVTQTVDGGTGNDYLYI 244

Qy        614 SKDGGFGNITVDGTSATEAGSYTVNRKVARGDIYHEVVKRQETKVGKRTETIQYRDYELR 673
Db        245 DYSNAIASITSTYNPTTNQGLITAD----------------------TNQVRYKNIERL 281

Qy        674 KVGYGYQSTDNLKSVEEVIGSQFNDVFKGSKFNDIFHSGE-GDDLLDGGAGDDRL----- 727
Db        282 ----------------EVKGTAYDDNIVGSNGDDTLNGGNGGNDTLKGGAGNDILRINDN 325

Qy        728 -------FGGKGNDRLSGDE--GDDLLDGGSGDD----------------VLNGGAGNDV 762
Db        326 SRNTNTAYGGTGNDFLYAEATGGTSLLDGGDGNDYLSIANSYSSVFGRSNTLNGGVGNDT 385

Qy        763 YIFRKGDGNDTLYDGTGNDKLAFAD----ANISD------------IMIERTKEGIIVKR 806
Db        386 LNAKATYGGNFLNGGVGND-FIYADYSSGTNILDGGDRDDYLSAAGASFHFDNTFFTLKG 444

Qy        807 NDHSGSINIPRWYITSNLQNYQSN-----KTDHKIEQLIGKDGS-YITSDQ 851
Db        445 GTGNDTLNANSSFSTNNLDGGTGNDRLTVENSYKNNLLNGGDGNDYLKGGQ 495
```

RESULT 11
US-09-813-214A-9
; Sequence 9, Application US/09813214A
; Patent No. US20020177200A1
; GENERAL INFORMATION:
; APPLICANT: Tucker, Kenneth
; APPLICANT: Plosila, Laura
; TITLE OF INVENTION: MORAXELLA CATARRHALIS OUTER MEMBRANE PROTEIN-106 POLYPEPTIDE, GENE
; TITLE OF INVENTION: SEQUENCE AND USES THEREOF
; FILE REFERENCE: 7969-089-999
; CURRENT APPLICATION NUMBER: US/09/813,214A
; CURRENT FILING DATE: 2000-03-20
; PRIOR APPLICATION NUMBER: 08/968,685
; PRIOR FILING DATE: 1997-11-12
; NUMBER OF SEQ ID NOS: 20

```
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 9
;   LENGTH: 2122
;   TYPE: PRT
;   ORGANISM: Moraxella catarrhalis
US-09-813-214A-9

  Query Match              4.9%;  Score 226.5;  DB 10;  Length 2122;
  Best Local Similarity   20.3%;  Pred. No. 6.9e-07;
  Matches  227;  Conservative  130;  Mismatches  362;  Indels  397;  Gaps   54;

Qy         14 GLNSTKSGLKNLYLAIPKDYDPQKGGTLNDFIKAADELGIARLAEEPNHTETAKK----S 69
Db        639 GINAGGKELTNVQSAI----NPATNGGQLDFM---------------NRLSTANTEKSGS 679

Qy         70 VDTVNQFLSLTQTGIAISATKLEKFLQKHSTNKLAKGLDSVENIDRKLGKASNVLSTLSS 129
Db        680 AATIKDLYNLSQVPLTFAG-------------------DTGPNVTKKLGEILKV------ 714

Qy        130 FLGTALAGIELDSLIKKGDAAPDALAKASIDLI-----NEIIGNLSQSTQTIEAFSSQLA 184
Db        715 ---------------KGGKTTADDLTKNNIGVVADSTDNSLTVKLAKTLSDLDAVNTKTL 759

Qy        185 KLGSTISQAKGFSNIGNKLQN--LNFSKTNLGLEIIT--------GL---------LSGI 225
Db        760 TASDKVTVDSG--NNTAKLQNGDLTFSKQNTGATPATNSKTIGVDGLKFTDNNGIALDGT 817

Qy        226 S------AGFALADKNASTGK------KVAAG-FELSNQVIGNVTKAISSYVLAQRVAAG 272
Db        818 TYITKDKVGFAKQDGSLDKSKPYLDKDKLKVGEVEITTNGINAGGKAIT----------G 867

Qy        273 LSTTGAVAALITSSIMLAISPLAFMNAADK-------------FNHANALDEFAKQF--- 316
Db        868 LSNTLTDATNATTG---HVTQLGIVDSTDKTRAASIGDVLNAGFNLKNNGD--AKDFVST 922

Qy        317 -------------RKFGYDGDHLLAEYQRGV-GT---------------IEASLTTISTA 347
Db        923 YDTVDFINGNATTAKVTYDGKASKVAYDVNVDGTTIHLTGADGNKNQIGVKTTTLTKTDA 982

Qy        348 LG------AVSAGVSAAAVGSAVGTPIALLVAGVTGLISGILEASKQAMFESVANRLQGK 401
Db        983 KGDKAINFSVNSGDDKALIN-------AKDIADNLNTLAGEIRNTKGTADTALQTFQVKK 1035

Qy        402 ILEWEKQNGGQNYFDKGYDSRYAAYLANNLKFLSELNKELEAERVIAITQQRWDNNIGEL 461
Db       1036 VKENGDDDNDADTITVGKDAK--TNQVNTLKLKGKNGLDIQTNK---------------- 1077

Qy        462 AGITKLGERIKSGKAYADAFEDGKKVEAGSNITLDAKTGIIDISNSNGKKTQALHFTSPL 521
Db       1078 DGTVTFGINTQSG------------LKAGNNTTL----------NNNGLSIKN------- 1108

Qy        522 LTAGTESRERLTNGKYSYINKLKFGRVKNWQVTDGEASSKLDFSKVIQRVAETEGTDEIG 581
Db       1109 -TAGNEQIQVGADG-------VKFAKVNN-----GVVGAGIDGTTRITR-------DEIG 1148

Qy        582 LIVNAKAGNDDIFVGQGKMNIDGGDGHDRVFYSKDGGFGNITVDGTSATEAGSYTV---- 637
Db       1149 F-----AGTNG--------SLDKSKPH----LSKDG----INAGGKKITNIQSGEIAQNS 1187

Qy        638 NRKVARGDIYHEVVKRQETKVGKRTETIQYRDYELR---------KVGYGYQSTDNLKSV 688
Db       1188 NDAVTGGKIY-DLKTELENKISSTAKTAQNSLHEFSVADEQGNNFTVSNPYSSYDTSKTS 1246

Qy        689 EEV-------IGSQFN--------DVFKG---SKFNDIFHSGEGDDLLDGGAGDDRLFG- 729
Db       1247 DVITFAGENGITTKVNKGVVRVGIDQTKGLTTPKLTVGNNNGKG-IVIDSQNGQNTITGL 1305

Qy        730 ---------GKGNDRLSGDEG-----DDLLDGGSGDDVLNGG---AGN----------DV 762
Db       1306 SNTLANVTNDKGSVRTT-EQGKIIKDEDKTRAASIVDVLSAGFNLQGNGEAVDFVSTYDT 1364

Qy        763 YIFRKGDGNDT----LYDGTGNDKLAFADANISDIMIE--RTKEGI-------------- 802
Db       1365 VNF--ADGNATTAKVTYDDTSKTSKVVYDVNVDDTTIEVKDKKLGVKTTTLTSTGTGANK 1422

Qy        803 -----------IVKRNDHSGSINIPRWYITSNLQNYQSNKTDHKIEQLIGKDGSYITSDQ 851
Db       1423 FALSNQATGDALVKASDIVAHLNT----LSGDIQTAKGASQANSSAGYVDADGNKVIYDS 1478

Qy        852 ID-KILQDKKDGTVITSQELKK--------------------------LADENKSQKLS 883
Db       1479 TDNKYYQAKNDGTVDKTKEVAKDKLVAQAQTPDGTLAQMNVKSVINKEQVNDANKKQGIN 1538

Qy        884 ASDI-------ASSLNKLVGSMALFGTANSVSSNAL 912
Db       1539 EDNAFVKGLEKAASDNKTKNAAVTVGDLNAVAQTPL 1574

RESULT 12
US-09-932-183A-2
; Sequence 2, Application US/09932183A
; Patent No. US20020127641A1
; GENERAL INFORMATION:
;  APPLICANT: Estell, David A.
;  TITLE OF INVENTION: Proteases From Gram-Positive Organisms
;  FILE REFERENCE: GC394C1-US
;  CURRENT APPLICATION NUMBER: US/09/932,183A
;  CURRENT FILING DATE:  2001-08-17
;  PRIOR APPLICATION NUMBER: US 09/308,375
;  PRIOR FILING DATE: 1999-05-14
;  PRIOR APPLICATION NUMBER: PCT/US98/18828
;  PRIOR FILING DATE: 1998-09-08
;  PRIOR APPLICATION NUMBER: EP9719636.4
;  PRIOR FILING DATE: 1997-09-15
;  NUMBER OF SEQ ID NOS: 3
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 2
;   LENGTH: 2285
;   TYPE: PRT
;   ORGANISM: Bacillius subtilis
US-09-932-183A-2

  Query Match              4.7%;  Score 219;  DB 10;  Length 2285;
  Best Local Similarity   20.6%;  Pred. No. 2.7e-06;
  Matches  225;  Conservative  153;  Mismatches  375;  Indels  340;  Gaps   55;

Qy          4 INVIKSNIQAGLNSTKSGLKNLYLAIPKDYDPQKGGTLNDFIKAADELGIARLAEEPNHT 63
Db        366 IDTLMTNIRRVMNEPDYKYNEL-LQESIDLGDTLSNKITDILQMTGDFG--RMGFDESEL 422

Qy         64 ETAKKSVDTVNQFLSLT--QTGIAISATKLEKFLQKHSTNKLAKGLDSVEN------ID- 114
Db        423 STLTKTAQVLQNVSDLTPDDTVNTLTAAMLNFNIAANDSISIADKLNEVDNNYAVTTLDL 482

Qy        115 ----RKLGKASNVLSTLSSFLGTALAGIELDSLIKKGDAAPDALAKASIDLINEIIGNLS 170
Db        483 ANSIRKAG------STASTF------GVELNDLI----GYTTAIASTTRESGN-IVGN-- 523

Qy        171 QSTQTIEAFSSQLAKLGSTISQAKGFSNIGNKLQNLNFSKTNLG-LEIITGLLSGISAGF 229
Db        524 -SLKTI------FARIGNNQSSIKALEQIGISV------KTAGGEAKSASDLISEVAGKW 570

Qy        230 -ALAD---KNASTGKKVAAGFELS--NQVIGNVTKAISSYVLAQRVA-AGLSTTGAV--- 279
Db        571 DTLSDAQKQNTSIG--VAGIYQLSRFNAMMNN-------FSIAQNAAKTAANSTGSAWSE 621

Qy        280 ---------------------AALITSSIMLAISPLAFMNAADKFNHANALDEFAKQFRK 318
Db        622 QQKYADSLQARVNKLQNNFTEFAIAASDAFISDGLIEFTQAA-----GSLLNASTGVIKS 676

Qy        319 FGYDGDHLLAE------YQRGVGTIEASLTTISTALG-------AVSAGVSAAAVGSAV- 364
Db        677 VGFLPPLLAAVSTATLLLSKNTRTLASSLILGTRAMGQETLATAGLEAGMTRAAVASRVL 736

Qy        365 -----GTPIALLVAG--------VTGLISGILEA---------SKQAMFESVANRLQG-- 400
```

```
Db        737 KTALRGLLVSTLVGGAFAALGWALESLISSFAEAKKAKDDFEQSQQTNVEAITTNKDSTD 796

Qy        401 KILEWEKQ-----------NGGQNYFD------------KGYDSRYAAYLANNLKFLS 435
Db        797 KLIQQYKELQKVKESRSLTSDEEQEYLQVTQQLAQTFPALVKGYDSQGNAILKTN----K 852

Qy        436 ELNKELEAERVIAITQQRWDNNIGELAGITKLGERIKSGKAYADAFEDGKKVEAGSNITL 495
Db        853 ELEKAIE-------------NTKEYLALKKQETRDSAKKTFEDASKEIKK--------- 889

Qy        496 DAKTGIIDISNSNGKKTQALHFTSPLLTAGTESRERLTNGKY--SYINKLKFGRVKNWQV 553
Db        890 --------------------------------SKDELKQYKQIADYNDK---GRPK-WDL 913

Qy        554 TDGEASSKLDFSKVIQRVAETEGTDEIGLIVNAKAGNDDIFVGQGKMNIDGGDGHDRVFY 613
Db        914 IADDDDYKVAADKAKQSMLKAQSDIESG---NAKVKDSVLSIANAYSSID--------- 960

Qy        614 SKDGGFGNITVDGTSATEAGSYTVNRKVARGDIYHEVVKRQETKVGKRTETIQYRDYELR 673
Db        961 ----------ISNTLKTSI-SDVVNKLNLKDDLDPEELEKFSSSLGKLQEKMQ------K 1003

Qy        674 KVGYGYQST-DN----LKSVEEVIGSQFN--DVFK-----------GSKFNDIFHSGEG 714
Db       1004 ALDSGDEKAFDNAKKDLQSLLETYSKSDSSIDVFKMSFDKAQKNIKDGDKSLSSVKSEVG 1063

Qy        715 D--DLLDGGAGDDRLFGGKGNDRLSGDEGDDL----------LDGGSGDDVLNGGAGNDV 762
Db       1064 DLGETLAEAGNEAEDFGKKLKEALDANSVDDIKAAIKEMSDAMQFDSVQDVLNG----DI 1119

Qy        763 YIFRKGDG---NDTLYDGTGNDKLA---FADANISDIMIERTKEGIIVKRNDHSGSINIP 816
Db       1120 FNNTKDQVAPLNDLL------EKMAEGKSISANEANTLIQKDKE-LAQAISIENGVVKIN 1172

Qy        817 RWYITSNLQNYQSNKTDHKIEQLIGKDGSYITSDQI---DKILQDKKDGTVIT------- 866
Db       1173 R---------------DEVIKQRKVKLDAY--NDMVTYSNKLMKTEVNNAIKTLNADTLR 1215

Qy        867 SQELKKLADENKSQKLSA------------SDIASSLNKLVGSMALFGTANSVSSNALQ 913
Db       1216 IDSLKKLRKERKLDMSEAELSDLEVKSINNVADAKKELKKLEEKMLQPGGYSNSQIEAMQ 1275

Qy        914 PITQPTQGILAPS 926
Db       1276 SVKSALESYISAS 1288
```

RESULT 13
US-10-192-584-7
; Sequence 7, Application US/10192584
; Publication No. US20030027987A1
;   GENERAL INFORMATION:
;   APPLICANT: TOKUNAGA, Eiji
;   SAKAGUCHI, Masashi
;   MATSUO, Kazuo
;   HAMADA, Fukusaburo
;   TOKIYOSHI, Sachio
;   TITLE OF INVENTION: NOVEL POLYPEPTIDE FROM HAEMOPHILUS
;   PARAGALLINARUM AND PROCESS FOR PREPARING THE SAME
;   NUMBER OF SEQUENCES: 8
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: BROWDY AND NEIMARK
;   STREET: 624 Ninth Street, N.W., Suite 300
;   CITY: Washington
;   STATE: D.C.
;   COUNTRY: USA
;   ZIP: 20001
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: Patent In Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/10/192,584
;   FILING DATE: 11-Jul-2002
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US/09/077,098
;   FILING DATE: 19-May-1998
;   APPLICATION NUMBER: PCT/JP97/03222
;   FILING DATE: 12-SEP-1997
;   APPLICATION NUMBER: JP 27,148/1996
;   FILING DATE: 19-SEP-1996
;   ATTORNEY/AGENT INFORMATION:
;   NAME: KORNBAU, Anne M.
;   REGISTRATION NUMBER: 25,618
;   REFERENCE/DOCKET NUMBER: TOKUNAGA=1
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 202-628-5197
;   TELEFAX: 202-737-3528
;   INFORMATION FOR SEQ ID NO: 7:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 2039 amino acids
;   TYPE: amino acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;   MOLECULE TYPE: peptide
;   SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-10-192-584-7

```
  Query Match             4.6%;  Score 216;  DB 15;  Length 2039;
  Best Local Similarity   20.2%;  Pred. No. 3.7e-06;
  Matches  215;  Conservative  126;  Mismatches  424;  Indels  298;  Gaps   44;

Qy         11 IQAGLNSTKSGLKNLYLAIPKDYDPQKGGTLNDFIKAADELGIARLAEEP--------- 60
Db        337 IAAGEMSTAVGSLAFALA-----DKSTAMGLRSFV-AKDAVGGTAIGEESRTFAKDSVAI 390

Qy         61 -NHTETAKKSVDTVNQFLSLTQTG-IAISA----------TKLEKFLQKHSTNKLAKGLD 108
Db        391 GNKTEASNAGSMAYGYKAKAVGAGAIAIGAEVAAGAEFDSSQAGNLLLNRGAYATLKSAD 450

Qy        109 SVENIDRKLGKASNVLSTL---------------SSFLGTALAGIELDSLIKKGDAAPDA 153
Db        451 KSDDI--KAGDAINVFTQFFDNMLTQGSHLTYENTTYLTTSAGDIK-KTLAAVGDGGKNA 507

Qy        154 LAKASIDLINEIIGN--LSQSTQTIEAFSSQLAK------LGS-----------TISQAK 194
Db        508 IA---------IGNKTFASKANSVALGSYALASAQNAFALGSYSLVSPLAANTIVIGVG 557

Qy        195 GFSNIGNKLQNLNFSKTNLGLEIITGLLSGISAGFALADKNASTGKKVAAGFELSNQVIG 254
Db        558 GYATGSNSFVGGSWVSTLSARTVVLGYSASIS-----SDSHDSLAMGV-------NAFIG 605

Qy        255 NVTKAISSYVLAQRVAAGLSTTGAVAALITSSIMLAISPLAFMNAADKFNHA----NALD 310
Db        606 NGSNS----------SLALGTGSTIAKNAKSPDSLAIGKDSRIDAKDTDNGVLYTPQVYD 655

Qy        311 EFAKQFRKFGYDGDHLL--------AEYQRGVGTIEASLTTI------------STALGA 350
Db        656 ETTRAFRTFDENKDYMRQAMALGFNAKVSRGKGKMETGINSMAIGARSQATLQNSTALGV 715

Qy        351 ------------VSAGVSAAAVGSAVGTPIALLVAG-------VTGLISGIL--EASKQA 389
Db        716 NAKTDYTWEQLEADPWVSKGAISIPTSGKIGVISVGSKGSERRIVNVASGSLDTDAVNVA 775

Qy        390 MFESVANRLQGKILEWEKQNGGQNYFD------------------KGYDSRYAAYLANNL 431
Db        776 QLKTIEERFQSEI-DLLQNGGGVQYLSVEKTNINGEAGRVASQIRKGESYKRYVKLKTQL 834

Qy        432 KFLSELNKELEAERVIAITQQRWDNNIGELAGITKLGERIKSGKAYADAFEDGKKVEAGS 491
Db        835 LYL-DARKKLNGEKFDQTSLDKISKAVQEL-------EAEYSGELKTTASELNRVA---- 882

Qy        492 NITLDAKTGIIDISNSNGKKTQALHFT-----------SPLLTAGTESRERLTNGKYSYI 540
Db        883 -MQLNAETTVNDFGKFNQYKTQIENATNADSEKNVGGLSPQVIAQLKANNNYLNDGAKGQ 941
```

```
Qy        541 NKLKFGRVKNWQVTDGEASSKLDFSKVIQRVAETEGTDEIGLIVNAKAGN---------- 590
Db        942 DSIAFG----WQAKTSEANNGLAGKQAI----------AIGFQANSSAENAISIGTNSDT 987

Qy        591 ---DDIFVGQGKMNIDGGDGHDRVFYSKDGGFGNITVDGTSATEAGSYTVNR-------- 639
Db        988 SMTGAVAIGKGATVTAG--GKPSIALGQDSTVANSAISRTSSVMINGLTFNNFAGSPETL 1045

Qy        640 ---------------KVARGDIYHEVVKRQETKVGKRTETIQYRDYELRKVGYGYQSTDN 684
Db       1046 GVLSIGTAGKERKIVNVAAGDI------SQTSTEAINGSQLYATNFMLNKLA-------- 1091

Qy        685 LKSVEEVIGSQFNDVFKGS-KFNDIFHSGEGDDLLDGGAGDDRLFGGKGN-----DRLSG 738
Db       1092 -QSVKTNFGGNANLATDGTITFTNI-----------GGTGQDTIHDAINNVLTKLISLSA 1139

Qy        739 DEGDDLLDGGSGDDVLNG-----------GAGNDVYIFRKGDGNDTLYDGTGND-KLAFA 786
Db       1140 TEEEEVVSGEAVYDALKGAKPTVSAEANKGITGLVDVVKKANSPITVEPSTDNNKKKTFT 1199

Qy        787 DANISDIMIERTKEGIIVKRNDHSG-SINIPRWYITSNLQNYQSNKTDHKIEQLIGKDGS 845
Db       1200 VGLMKDI------EGVNSITFDKSGQDLNQVTGRMSSAGLTFKKGDTTNGSTTTFAEDGL 1253

Qy        846 YI---TSDQIDKILQDKKDGTVITSQELKKLADENK--SQKLS 883
Db       1254 TIDSTTNSAQTNLVKVSRDGF-----SVKNGSDESKLASTKLS 1291
```

```
RESULT 14
US-10-192-584-6
; Sequence 6, Application US/10192584
; Publication No. US20030027987A1
;   GENERAL INFORMATION:
;        APPLICANT: TOKUNAGA, Eiji
;                   SAKAGUCHI, Masashi
;                   MATSUO, Kazuo
;                   HAMADA, Fukusaburo
;                   TOKIYOSHI, Sachio
;        TITLE OF INVENTION: NOVEL POLYPEPTIDE FROM HAEMOPHILUS
;                            PARAGALLINARUM AND PROCESS FOR PREPARING THE SAME
;        NUMBER OF SEQUENCES: 8
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: BROWDY AND NEIMARK
;             STREET: 624 Ninth Street, N.W., Suite 300
;             CITY: Washington
;             STATE: D.C.
;             COUNTRY: USA
;             ZIP: 20001
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: Patent In Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/192,584
;             FILING DATE: 11-Jul-2002
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: US/09/077,098
;             FILING DATE: 19-May-1998
;             APPLICATION NUMBER: PCT/JP97/03222
;             FILING DATE: 12-SEP-1997
;             APPLICATION NUMBER: JP 27,148/1996
;             FILING DATE: 19-SEP-1996
;        ATTORNEY/AGENT INFORMATION:
;             NAME: KORNBAU, Anne M.
;             REGISTRATION NUMBER: 25,618
;             REFERENCE/DOCKET NUMBER: TOKUNAGA=1
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 202-628-5197
;             TELEFAX: 202-737-3528
;   INFORMATION FOR SEQ ID NO: 6:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 2042 amino acids
;             TYPE: amino acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        MOLECULE TYPE: peptide
;        SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-10-192-584-6

  Query Match             4.6%;  Score 211.5;  DB 15;  Length 2042;
  Best Local Similarity   22.0%;  Pred. No. 7.8e-06;
  Matches  225;  Conservative  123;  Mismatches  386;  Indels  289;  Gaps   53;

Qy         11 IQAGLNSTKSGLKNLYLAIPKDYDPQKGGTLNDFIKAADELGIARLAEEP---------- 60
Db        334 IAAGEMSTAVGSLAFALA-----DRSTAMGLRSFV-AKDAVGGTAIGEESRTFAKDSVAI 387

Qy         61 -NHTETAKKSVDTVNQFLSLTQTGIAISATKL---EKFLQKHSTNKLAKGLDSVENI--- 113
Db        388 GNKTEASNAGSMAYGYKAKAVGAGAIAIGTEVAAGAKF-NSHQTGNLLQDNNAYATLKNA 446

Qy        114 ----DRKLGKASNVLSTLSSFLGTALAGIELDSLIKKGDAAPDALAKASIDLINEIIGNL 169
Db        447 DKSDDTKTGNAITVFT--QSF----------DNMLTNG------------LPLVSENETYL 483

Qy        170 SQSTQTIEAFSSQLAKLGSTISQAKGFSN---IGNKLQNLNFSKTNLGLEIITGLLSGIS 226
Db        484 TTSAGAIK-------KTATTDSSAGGGKNAIAIGSK----TFASKANSVALGSYALADAQ 532

Qy        227 AGFALAD---------------------KNASTGKKVAAGFELSNQVIGNVTK------ 258
Db        533 NAFALGSYSFVESSATNTITIGVGSYAKGKNSFLGGTWASTLSDRTVVLGNSTSISSGSQ 592

Qy        259 ---AISSYVL-----AQRVAAGLSTTGAVAALITSSIMLAISPLAFMNAADKFN----HA 306
Db        593 NALAIGVNVFIGNDSASSLALGMGSTIAKSAKSPDS--LAIGKEARIDAKDTDNGTLYQP 650

Qy        307 NALDEFAKQFRKFGYDGDHLL---------AEYQRGVGTIEASLTTISTALGAVSAGVSAA 358
Db        651 QVYDETTRAFRNFNESSDYMRQAMALGFNAKVSRGVGKMETGIN--SMAIGAYAQATLQN 708

Qy        359 AVGSAVGT------------------PIALLVAGVTGLISGILEASKQAMFESVANRLQG 400
Db        709 STALGVGSKTDYTWEQLETDPWVSEGAISIPTSGKTGVISVGSKGSER----RIVNLASG 764

Qy        401 KILEWEKQNGGQNYFDKGYDSRYAAYLAN-NLKFLSELN--KELEAERVIAITQQRWDNN 457
Db        765 ---------------SSDTDAVNVAQLKTVEERFLSEINLLQNGGGVKYLSVEKTNINGQ 809

Qy        458 IGELAGITKLGE------RIKSGKAYADAFE--DGKKVEAGSNITLDAKTGIIDISNSNG 509
Db        810 SGRVASQIRKGENYERYVKLKTQLLYLDARGKLNGEKFDQNSLNKIRAVVQELEAEYSGE 869

Qy        510 KKTQALHFTSPLLTAGTESRERLTNGKYSYINKLKFGRVKNWQVTDGEASSKLDFSKVI- 568
Db        870 LKTTA----SALNQVATQLEQEVTTNNFDKFNQYK---------TQIENASNADSARNVG 916

Qy        569 ----QRVAE----------TEGTDEIGLIVNAK--------AGNDDIFVG-----QGKM 600
Db        917 GLTPQAIAQLKANNNYLNDGAKGQDSIAFGWQAKTSGANNGLAGKQAIAIGFQANSSAEN 976

Qy        601 NIDGGDGHD-----RVFYSKD-----GGFGNITV--DGTSATEAGSYTVNRKVARGDIYH 648
Db        977 AISIGTNSDTSMTGAVAIGKGATVTAGGKPSIALGQDSTVANSAISRT-SSPMINGLIFN 1035

Qy        649 EVVKRQETKVGKRTETIQYRDYELRKVGYGYQSTDNLKSVEEVIGSQ-FNDVFKGSKFND 707
Db       1036 NFAGSPET-LGVLSIGTAGRERKIVNVAAGDVSQ---ASTEAINGSQLYATNFMLSKVAQ 1091

Qy        708 IFHSGEGDDLLDGGAGDDRLFGGKGNDRLSGDEGDDLLD--GGSGD----DVLNGGAGND 761
Db       1092 SVKSN---------------FGGNVN---LGTDGTITFTNIGGTGQATIHDAINNVLTKG 1133
```

```
Qy        762 VYIFRKGDGNDTLYDGTGN--DKLAFADANISDIMIERTKEGIIVKRNDHSGSINIPRWY 819
Db       1134 IYL--KADQNDP----TGNQGQKVELGNA----ITLSATNQWA-------NNGVN----Y 1172

Qy        820 ITSNLQNYQSN--------KTDHKIEQLIGKDGSYITS------DQIDKILQDKK-DGTV 864
Db       1173 KTNNLTTYNSQNGTILFGMREDPSVKQITA--GTYNTTGDANNKNQLNNTLQQTTLEATG 1230

Qy        865 ITS 867
Db       1231 ITS 1233
```

RESULT 15
US-09-797-862-33
; Sequence 33, Application US/09797862
; Patent No. US20020102276A1
; GENERAL INFORMATION:
; APPLICANT: PEAK, IAN RICHARD ANSELM
; APPLICANT: JENNINGS, MICHAEL PAUL
; APPLICANT: MOXON, E. RICHARD
; TITLE OF INVENTION: NOVEL SURFACE ANTIGEN
; FILE REFERENCE: 065064/0134
; CURRENT APPLICATION NUMBER: US/09/797,862
; CURRENT FILING DATE: 2001-05-03
; PRIOR APPLICATION NUMBER: PCT/AU98/01031
; PRIOR FILING DATE: 1998-12-14
; PRIOR APPLICATION NUMBER: GB 9726398.2
; PRIOR FILING DATE: 1997-12-12
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 33
; LENGTH: 2353
; TYPE: PRT
; ORGANISM: Haemophilus influenzae
US-09-797-862-33

```
  Query Match             4.3%;  Score 202;  DB 10;  Length 2353;
  Best Local Similarity   20.8%;  Pred. No. 4.7e-05;
  Matches  212;  Conservative  129;  Mismatches  338;  Indels  338;  Gaps   48;

Qy         62 HTETAKKSVDTVNQFLSLTQTGIAISATKLEKFLQKHSTNKLAKGLDSVENIDRKLGKAS 121
Db         26 HTKRASATVET-----------AVLATLLFATVQANAT-------DEDEELD-PVVRTA 65

Qy        122 NVLSTLSSFLGTALA--------GIELDSLIKKGDAAPDALA-KASIDLINEIIGNLSQS 172
Db         66 PVLSFHSDKEGTGEKEVTENSNWGIYFDN---KGVLKAGAITLKAGDNL--KIKQNTDES 120

Qy        173 TQTIEAFSSQLAK----LGSTISQAKGFSNIGNKL------QNLNFSKTNLGLEIITGLL 222
Db        121 TNA-SSFTYSLKKDLTDLTSVATEKLSFGANGDKVDITSDANGLKLAKTGNGNVHLNGLD 179

Qy        223 SGISAGFALADKNASTGKKVAAGFELSNQVIGNVTKAISSYVLAQRVAAGLSTTGAVAA- 281
Db        180 S------TLPDAVTNTGVLSSSSF------TPNDVEKTRAATVKDVLNAGWNIKGAKTAG 227

Qy        282 --LITSSIMLAISPLAFMNAADKFNHANALDEF--AKQFRKF------------GYDGD 324
Db        228 GNVESVDLVSAYNNVEFI-TGDK----NTLDVVLTAKENGKTTEVKFTPKTSVIKEKDGK 282

Qy        325 HLLAEYQRGVGTIEASLTTISTALGAVSAGVSAAAVGSAV---GTPIALLVA-GVTGLIS 380
Db        283 LFTGKENNDTNKVTSNTATDNTDEG--NGLVTAKAVIDAVNKAGWRVKTTANGQNGDFA 340

Qy        381 GILEASKQAMFESVANRLQGKILEWEKQNGGQNYFDKGYDSRYAAYLANNLKF------- 433
Db        341 TVASGT-NVTFESG----DGTTASVTKDTNG-----NGITVKYDAKVGDGLKFDSDKKIV 390

Qy        434 ------------LSELNKELEAERVI-------AITQQRWD-----NNIGELAGITKLGE 469
Db        391 ADTTALTVTGGKVAEIAKEDDKKKLVNAGDLVTALGNLSWKAKAEADTDGALEGISK-DQ 449

Qy        470 RIKSGKAYADAFEDGKKVEA---GSNITL---DAKTGIIDIS----------------- 505
Db        450 EVKAGETV--TFKAGKNLKVKQDGANFTYSLQDALTGLTSITLGGTTNGGNDAKTVINKD 507

Qy        506 --------NSNGKKTQALHFTSPLLTAGTESRERLTNGKYSY----------------- 539
Db        508 GLTITPAGNGGTTGTNTISVTKDGIKAGNKAITNVASGLRAYDDANFDVLNNSATDLNRH 567

Qy        540 --------INKLKFGRVKNWQVTDGEASSKLDFSKVIQRVAETEGTDEIGLIVNAKAGND 591
Db        568 VEDAYKGLLNLNEKNANKQPLVTDSTAATVGDLRKLGWVVSTKNGTKEESNQV--KQADE 625

Qy        592 DIFVGQGKMNIDGGDGHDRVFYSKDGGFGNITVDGTSATEAGSYTVNRKVARGDIYHEVV 651
Db        626 VLFTGAGAATV----------------------TSKSENGKHTITVSVA--------- 652

Qy        652 KRQETKVGKRTETIQYRDYELRKVGYGYQSTDNLKSVEEVIGSQFNDVFKGSKFNDIFHS 711
Db        653 ---ETKADCGLE----KDGDTIKLKVDNQNTDNVLTV----GNNGTAVTKG--------- 692

Qy        712 GEGDDLLDGGAGD-DRLFGGKGNDRLSGDEGDDLLDGGSGDDVLNGGAGNDVYIFRKGDG 770
Db        693 --GFETVKTGATDADR-------------------------------------------G 707

Qy        771 NDTLYDGTGNDKLAFAD---ANISDIMIERTKEGIIVKRNDHSGSINIPRWYITSNLQNY 827
Db        708 KVTVKDATAND----ADKKVATVKDVATAINSAATFVKTENLTTSID------------ 750

Qy        828 QSNKTDHKIEQLIGKDGSYITSDQI----DKILQDKKDGTVIT-----SQELK--KLAD- 875
Db        751 EDNPTDN------GKDDALKAGDTLTFKAGKNLKVKRDGKNITFDLAKNLEVKTAKVSDT 804

Qy        876 -----------ENKSQKLSASDIASSLNKLVGSMALFGTANSVSSNALQPITQPTQG 921
Db        805 LTIGGNTPTGGTTATPKVNITSTADGLNFAKETADASGSKNVYLKGIATTLTEPSAG 861
```

RESULT 16
US-10-369-493-19032
; Sequence 19032, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 19032
; LENGTH: 239
; TYPE: PRT
; ORGANISM: Anabaena PCC7120
US-10-369-493-19032

```
  Query Match             4.3%;  Score 200;  DB 12;  Length 239;
  Best Local Similarity   38.3%;  Pred. No. 2e-06;
  Matches   49;  Conservative   19;  Mismatches   32;  Indels   28;  Gaps    6;

Qy        677 YGYQSTDNLKSVEEVIGSQFNDVFKGSKFNDIFHSGEGDDL---------LDGGAGDDRL 727
Db          5 HGTNSNDYL------YGGDESDTIKGFAGYDYLYGGLGDDYLYGGLDNDKLYGGLGNDKL 58

Qy        728 FGGKGNDRLSGDE---------GDDLLDGG--SGDDVLNGGAGND-VYIFRKGDGNDTLY 775
Db         59 YGGEGNDNLNGEQGNNSLYGEAGDDTLDAGNSTGNNLLDGGEGNDEIYVY-SSSGNNTLK 117
```

```
Qy        776 DGTGNDKL 783
Db        118 GGVGNDIL 125
```

RESULT 17
US-09-841-786-1
; Sequence 1, Application US/09841786
; Patent No. US20020054883A1
; GENERAL INFORMATION:
; APPLICANT: NAGARAJA, T. G.
; APPLICANT: STEWART, GEORGE C.
; APPLICANT: NARAYANAN, SANJEEV K.
; APPLICANT: CHENGAPPA, M. M.
; TITLE OF INVENTION: RECOMBINANT FUSOBACTERIUM NECROPHORUM LEUKOTOXIN
; TITLE OF INVENTION: VACCINE AND PREPARATION THEREOF
; FILE REFERENCE: 30296
; CURRENT APPLICATION NUMBER: US/09/841,786
; CURRENT FILING DATE: 2001-04-24
; PRIOR APPLICATION NUMBER: 09/558,257
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 3241
; TYPE: PRT
; ORGANISM: Fusobacterium necrophorum
US-09-841-786-1

```
  Query Match             4.3%;  Score 199;  DB 9;  Length 3241;
  Best Local Similarity   21.1%;  Pred. No. 0.00012;
  Matches  194;  Conservative  132;  Mismatches  362;  Indels  230;  Gaps   44;

Qy          5 NVIKSNIQAGLNSTKSG----LKNLYLAIPKDYDPQKGGTLNDFIKAADELGIARLAEEP 60
Db       2382 NIKQTAANAGIGGAAIGANVLVNNFGTAVEDRKNSEGKGT--EVLKTLDEV--------- 2430

Qy         61 NHTETAKKSVDTVNQFLSLTQTGIAISATKLEK---FLQKHSTNKLAKGLDSV-ENIDRK 116
Db       2431 -NKEQDKKVNDATKKILQ--SAGISTEDTSVKADRGDTQGEGIKAIVKTSDIIGKNVDIT 2487

Qy        117 LGKASNVLST-------LSSFLGT-ALAGIELDSLIKKGDAAPDALAKAS--IDLINEII 166
Db       2488 TEDKNNITSTGGLGTAGLASASGTVAVTNIKRNS----GVTVENSFVKAAEKVNVRSDIT 2543

Qy        167 GNLSQSTQTIEAFSSQLAKLGSTISQAK----GFSNI---GNKLQNLNF-----SKTNLG 214
Db       2544 GNVA-----LTAYQGPVGALGIGAAYAELNSNGRSNISIKNSKLLGKNIDVIVKDKSELR 2598

Qy        215 LEIITGLLSGISAGFALADK------------------NASTGKKVAAGFELSNQVIGN 255
Db       2599 AE-AKGLTVGAVAAGAIISKAKNEMNSEVEIEKSIFNEENRVTSPSKGIGREI------N 2651

Qy        256 VTKAISSYVLAQRVAAGLSTTGAVA-ALITSSIMLAISPLAFMNAADKFN----HANALD 310
Db       2652 VKVEKENRVTAESQGA---SVGAVAGAGIISEAKDAGS--SYLKVSTKSGRSIFHADNVN 2706

Qy        311 EFAKQFRKFGYDGDHLLAEYQRGVGTIEASLTTISTALGAVSAG-----VSAAAVGSAVG 365
Db       2707 MEATHKMKVTAVSKAVTGSVLGGVGVTKAEATAAGKTMVEVEEGNLFRTNRLNAISKVEG 2766

Qy        366 ------TPIALLVAGVTGLI--SGILEASKQAMFESVANRLQGKILEWEKQNGGQNYFDK 417
Db       2767 LDEDKVTAKSSVVSGNGGGIAGAGVNTSTAQSNTESVV-RLR-------KQDYENNDYTK 2818

Qy        418 GYDSRYAAYLANNLKFLSELNKELEAERVIAITQQRWDNNIGEL--AGITKLGERIKSGK 475
Db       2819 --------------KYISEVN-------ALALNDTKNEANIESLAVAGVHAQG----TNK 2853

Qy        476 AYADAFEDGKKVEAGSNIT-LDAKTGIIDISNSNGKKTQALHFTSPLLTAGTESRERLTN 534
Db       2854 AFTRSNKLTSTTVNGGNVSQLRAKALAKNENYGNVKGT-----GGALVGAETAAVENYTK 2908

Qy        535 GKYSYINKLKFGRVKNWQVTDGEASSKLDFSKVIQRVAETEGTDEIGLIV---------- 584
Db       2909 ---STTGALVAG---NWEIGD-----KLETIARDNTIVRVNGDGTKGGLVGKNGISVKNT 2957

Qy        585 ---NAKAGNDD--IFVGQGKMNIDGGDGHDRVFYSKDGGFGNITVDGTSATEAGSYTVNR 639
Db       2958 ISGETKSSIEDKARIVGTGSVNVDALNELDVDLQGKSGGYGGIGIGNV------------ 3005

Qy        640 KVARGDIYHEVVKRQETKVGKRTETIQYRDYELRKVGYGYQSTDNLKSVEEVIGSQFNDV 699
Db       3006 -----DVNNVIKKNVEAKIGRH------------------------AIVETTGKQEYQA 3035

Qy        700 FKGSKFNDIFHSGEGDDLLDGGAGDDRLFGGKGNDRLSGD-EGDDLLDGGSGDDVLNGGA 758
Db       3036 FTRAKVNIL-----------GKGDAAAAAAISNVHISNEMDIKNLAKQYASSQLITKNS 3083

Qy        759 GNDVYIFRKGDGNDTLYDGTGNDKLAFADANIS-DIMIERTKEGIIVKRNDHSGSINIPR 817
Db       3084 KNNITLASSSESNVNVH-GVAEARGAGAKATVSVKNQINRTNNVDLAGKIKTEGNINVYA 3142

Qy        818 WYITSNLQNYQSNKTDHK 835
Db       3143 GYD----KNYNISKTNSK 3156
```

RESULT 18
US-10-175-275-4
; Sequence 4, Application US/10175275
; Publication No. US20030171254A1
; GENERAL INFORMATION:
; APPLICANT: SASAKI, Ken
; APPLICANT: HARKNESS, Robin E.
; APPLICANT: LOOSMORE, Sheena M.
; APPLICANT: CHONG, Pele
; APPLICANT: KLEIN, Michel H.
; TITLE OF INVENTION: HIGH MOLECULAR WEIGHT MAJOR OUTER MEMBRANE PROTEIN OF
; TITLE OF INVENTION: MORAXELLA
; FILE REFERENCE: 1038-1235 MIS
; CURRENT APPLICATION NUMBER: US/10/175,275
; CURRENT FILING DATE: 2002-06-20
; PRIOR APPLICATION NUMBER: 08/945,567
; PRIOR FILING DATE: 1998-03-19
; PRIOR APPLICATION NUMBER: 08/431,718
; PRIOR FILING DATE: 1995-05-01
; PRIOR APPLICATION NUMBER: 08/478,370
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 08/621,944
; PRIOR FILING DATE: 1996-03-26
; PRIOR APPLICATION NUMBER: PCT/CA96/00264
; PRIOR FILING DATE: 1996-04-29
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 1833
; TYPE: PRT
; ORGANISM: Moraxella catarrhalis
US-10-175-275-4

```
  Query Match             4.2%;  Score 197;  DB 12;  Length 1833;
  Best Local Similarity   19.6%;  Pred. No. 7.3e-05;
  Matches  238;  Conservative  157;  Mismatches  407;  Indels  412;  Gaps   61;

Qy          2 SNINVIK----SNIQAGLNSTKSGLKNLYLAIPKDYDPQKGGTLNDFIKAADELGIARLA 57
Db        181 NNIGVVKEADNSGLKVKLAKTLNNLTEVNTTTLNATTTVKVGSSSS--TTAELLSDSLTF 238

Qy         58 EEPN-HTETAKKSVDTVN--QFLSLTQTGIAISATKLEKFLQKHSTNKLAKGLD------ 108
Db        239 TQPNTGSQSTSKTVYGVNGVKFTNNAETTAAIGTTRITR-------DKIGFARDGDVDEK 291

Qy        109 SVENIDRKLGKASNVLSTLSSFLGTALAGIELDSLIKKGDAAPDALAKASIDLINEIIGN 168
Db        292 QAPYLDKKQLKVGSVAITIDNGID---AGNKKISNLAKGSSANDAVT----------IEQ 338
```

```
Qy      169 LSQSTQTIEAFSSQLAKLGSTISQAKGFSNIGNKLQNLNFSKTNLGLEIITGLLSGISAG 228
            |  :  |:  |           |:  ||      ::   |  |:      |:|::         |  |
Db      339 LKAAKPTLNA--------GAGISVTPTEISVDAKSGNVTAPTYNIGVKTTELNSDGTSDK 390

Qy      229 FALADKNASTGKKVAAGFELS---NQVIGNVTKAISSYVL-------AQRVAAGLSTT-- 276
            |::   |  :  |     :        |:    |:|      |: |: :          |    :      ||
Db      391 FSV--KGSGTNNSLVTAEHLASYLNEVNRTADSALQSFTVKEEDDDDANAITVAKDTTKN 448

Qy      277 -GAVAAL--------------------------------------------ITSSIMLAI 291
             |||: |                                                      | :
Db      449 AGAVSILKLKGKNGLTVATKKDGTVTFGLSQDSGLTIGKSTLNNDGLTVKDTNEQIQVGA 508

Qy      292 SPLAFMNAADKFNHANALDEFAKQFR----KFGYDGDHLLAEYQRGVGTIEASLTTISTA 347
            : :  |  |         |  :|        |       |       |  |:  |               |  |  :   |
Db      509 NGIKFTNV----NGSNPGTGIANTARITRDKIGFAGS------DGAVDTNKPYLDQDKLQ 558

Qy      348 LGAV---SAGVSAAAVGSAVGTPIALLVAGVTGLISGILEASKQAMFESVANRLQGKILE 404
            :|  |      :  | ::|       |  |           : |::   :    |  :   |  :   :       :  |  :|  |
Db      559 VGNVKITNTGINAG--GKA--------ITGLSPTLPSIADQSSRNI--ELGNTIQDK--- 603

Qy      405 WEKQNGG--QNYFDKGYDSRYAAYLANN---LKFLSE------LNKELEAERVIAITQQR 453
             :|  |        :   :  |::            |  ||       :  |:|                 |               |   |    :
Db      604 -DKSNAASINDILNTGFN------LKNNNNPIDFVSTYDIVDFANGNATTATVTHDTANK 656

Qy      454 -----WDNNIGE----LAG---------ITKLGERIKSGKAYAD-----AFEDG----- 484
                 :|  |  :  :         |  |                      |||  :     :|         :          :  ||
Db      657 TSKVVYDVNVDDTTIHLTGTDDNKKLGVKTTKLNKTSANGNTATNFNVNSSDEDALVNAK 716

Qy      485 -----------------------------KKVEAGSN--------------------ITL 495
                                                     |||:   :|                                       :||
Db      717 DIAENLNTLAKEIHTTKGTADTALQTFTVKKVDENNNADDANAITVGQKNANNQVNTLTL 776

Qy      496 DAKTGIIDISNSNGKKTQALHFTSPLLTAGTESRERLTNGKYSYINK-----------L 543
              :  |:      ::  ||   |    ::  ||    |  ||          :      |  :|       |     |                      :
Db      777 KGENGLNIKTDKNGTVTFGINTTSG-LKAG---KSTLNDGGLSIKNPTGSEQIQVGADGV 832

Qy      544 KFGRVKNWQVTDGEASSKLDFSKVIQRVAETEGTDEIGLIVNAKAGNDDIFVGQGKMNID 603
            || :|  |          :|      :  |   :  |   :                    ||||                       ||         |
Db      833 KFAKVNN----NGVVGAGIDGTTRITR-------DEIG------------FTG-----TN 864

Qy      604 GGDGHDRVFYSKDGGFGNITVDGTSAT---EAGSYTVNR--KVARGDIYHEVVKRQETKVG 659
            |           :       ||||       |       |       |      ::|          |          |     |  ||     ::          |  |:
Db      865 GSLDKSKPHLSKDG----INAGGKKITNIQSGEIAQNSHDAVTGGKIY-DLKTELENKIS 919

Qy      660 KRTETIQYRDYELR---------KVGYGYQSTDNLKSVEEV-------IGSQFN------ 697
              :|  |      :|                             |        |    |    |     |:  :  :                 |   :  :  |
Db      920 STAKTAQNSLHEFSVADEQGNNFTVSNPYSSYDTSKTSDVITFAGENGITTKVNKGVVRV 979

Qy      698 --DVFKG---SKFNDIFHSGEGDDLLDGGAGDDRLFG----------GKGNDRLSGDEGD 742
              |    ||       |             ::|:|    ::|     |   :  :     |                      ||:  |   :  ::|:
Db      980 GIDQTKGLTTPKLTVGNNNGKG-IVIDSQNGQNTITGLSNTLANVTNDKGSVRTT-EQGN 1037

Qy      743 DLLD-----GGSGDDVLNGG---AGN----------DVYIFRKGDGNDT----LYDGTGN 780
             :  |                |||:  |           ||              |              |||   |            ||    |
Db     1038 IIKDEDKTRAASIVDVLSAGFNLQGNGEAVDFVSTYDTVNF--ADGNATTAKVTYDDTSK 1095

Qy      781 DKLAFADANISDIMIE--RTKEGI-------------------------IVKRNDHSGSI 813
                     |  |:   |    ||        |  |:                                                  :||   :|       .:
Db     1096 TSKVVYDVNVDDTTIEVKDKKLGVKTTTLTSTGTGANKFALSNQATGDALVKASDIVAHL 1155

Qy      814 NIPRWYITSNLQNYQSNKTDHKIEQLIGKDGSYITSDQID-KILQDKKDGTVITSQELKK 872
            |          :: ::|    :           :           :     ||:  :    |     |     |     |     ||||     ::|:  |
Db     1156 NT----LSGDIQTAKGASQANNSAGYVDADGNKVIYDSTDNKYYQAKNDGTVDKTKEVAK 1211

Qy      873 ----------------------------LADENKSQKLSASDI-------ASSLNKLVGSM 898
                                                   :   |   ||   |   ::      :                 :|   ||          :
Db     1212 DKLVAQAQTPDGTLAQMNVKSVINKEQVNDANKKQGINEDNAFVKGLEKAASDNKTKNAA 1271

Qy      899 ALFGTANSVSSNAL 912
                    |    |:|:        |
Db     1272 VTVGDLNAVAQTPL 1285
```

```
RESULT 19
US-10-175-282-4
; Sequence 4, Application US/10175282
; Publication No. US20030170657A1
; GENERAL INFORMATION:
;  APPLICANT: SASAKI, Ken
;  APPLICANT:  HARKNESS, Robin E.
;  APPLICANT:  LOOSMORE, Sheena M.
;  APPLICANT:  CHONG, Pele
;  APPLICANT:  KLEIN, Michel H.
;  TITLE OF INVENTION: HIGH MOLECULAR WEIGHT MAJOR OUTER MEMBRANE PROTEIN OF
;  TITLE OF INVENTION:  MORAXELLA
;  FILE REFERENCE: 1038-1234 MIS
;  CURRENT APPLICATION NUMBER: US/10/175,282
;  CURRENT FILING DATE:  2002-06-20
;  PRIOR APPLICATION NUMBER: 08/945,567
;  PRIOR FILING DATE: 1998-03-19
;  PRIOR APPLICATION NUMBER: 08/431,718
;  PRIOR FILING DATE: 1995-05-01
;  PRIOR APPLICATION NUMBER: 08/478,370
;  PRIOR FILING DATE: 1995-06-07
;  PRIOR APPLICATION NUMBER: 08/621,944
;  PRIOR FILING DATE: 1996-03-26
;  PRIOR APPLICATION NUMBER: PCT/CA96/00264
;  PRIOR FILING DATE: 1996-04-29
;  NUMBER OF SEQ ID NOS: 10
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
;   LENGTH: 1833
;   TYPE: PRT
;   ORGANISM: Moraxella catarrhalis
US-10-175-282-4

  Query Match             4.2%;  Score 197;  DB 12;  Length 1833;
  Best Local Similarity   19.6%;  Pred. No. 7.3e-05;
  Matches 238;  Conservative  157;  Mismatches  407;  Indels  412;  Gaps   61;

Qy        2 SNINVIK----SNIQAGLNSTKSGLKNLYLAIPKDYDPQKGGTLNDFIKAADELGIARLA 57
             :||  |:|       |  ::     |     |  :   |     :                  |  |:  :           |:  |     :
Db      181 NNIGVVKEADNSGLKVKLAKTLNNLTEVNTTTLNATTTVKVGSSSS--TTAELLSDSLTF 238

Qy       58 EEPN-HTETAKKSVDTVN--QFLSLTQTGIAISATKLEKFLQKHSTNKLAKGLD------ 108
             :||    :::     |:|      ||     :|    :      :|       ||      |::  :                :|:         |
Db      239 TQPNTGSQSTSKTVYGVNGVKFTNNAETTAAIGTTRITR-------DKIGFARDGDVDEK 291

Qy      109 SVENIDRKLGKASNVLSTLSSFLGTALAGIELDSLIKKGDAAPDALAKASIDLINEIIGN 168
                 :|:|     |     :|     |:  :   :         ||   :    |   :  ||  :|  ||:                                   |
Db      292 QAPYLDKKQLKVGSVAITIDNGID---AGNKKISNLAKGSSANDAVT----------IEQ 338

Qy      169 LSQSTQTIEAFSSQLAKLGSTISQAKGFSNIGNKLQNLNFSKTNLGLEIITGLLSGISAG 228
              |    :      |:  |           |:  ||      ::   |  |:      |:|::         |  |
Db      339 LKAAKPTLNA--------GAGISVTPTEISVDAKSGNVTAPTYNIGVKTTELNSDGTSDK 390

Qy      229 FALADKNASTGKKVAAGFELS---NQVIGNVTKAISSYVL-------AQRVAAGLSTT-- 276
            |::   |  :  |     :        |:    |:|      |: |: :          |    :      ||
Db      391 FSV--KGSGTNNSLVTAEHLASYLNEVNRTADSALQSFTVKEEDDDDANAITVAKDTTKN 448

Qy      277 -GAVAAL--------------------------------------------ITSSIMLAI 291
             |||: |                                                      | :
Db      449 AGAVSILKLKGKNGLTVATKKDGTVTFGLSQDSGLTIGKSTLNNDGLTVKDTNEQIQVGA 508

Qy      292 SPLAFMNAADKFNHANALDEFAKQFR----KFGYDGDHLLAEYQRGVGTIEASLTTISTA 347
            : :  |  |         |  :|        |       |       |  |:  |               |  |  :   |
Db      509 NGIKFTNV----NGSNPGTGIANTARITRDKIGFAGS------DGAVDTNKPYLDQDKLQ 558

Qy      348 LGAV---SAGVSAAAVGSAVGTPIALLVAGVTGLISGILEASKQAMFESVANRLQGKILE 404
            :|  |      :  | ::|       |  |           : |::   :    |  :   |  :   :       :  |  :|  |
Db      559 VGNVKITNTGINAG--GKA--------ITGLSPTLPSIADQSSRNI--ELGNTIQDK--- 603

Qy      405 WEKQNGG--QNYFDKGYDSRYAAYLANN---LKFLSE------LNKELEAERVIAITQQR 453
             :|  |        :   :  |::            |  ||       :  |:|                 |               |   |    :
Db      604 -DKSNAASINDILNTGFN------LKNNNNPIDFVSTYDIVDFANGNATTATVTHDTANK 656
```

```
Qy        454 -----WDNNIGE----LAG---------ITKLGERIKSGKAYAD-----AFEDG----- 484
                   :| |: :     ||          |||  :    :|      :     : ||
Db        657 TSKVVYDVNVDDTTIHLTGTDDNKKLGVKTTKLNKTSANGNTATNFNVNSSDEDALVNAK 716

Qy        485 ----------------------------KKVEAGSN-------------------ITL 495
                                          |||:  :|                   :||
Db        717 DIAENLNTLAKEIHTTKGTADTALQTFTVKKVDENNNADDANAITVGQKNANNQVNTLTL 776

Qy        496 DAKTGIIDISNSNGKKTQALHFTSPLLTAGTESRERLTNGKYSYINK-----------L 543
                : |:   :: ||  |  :: ||  | ||   :  | :|  |  |             :
Db        777 KGENGLNIKTDKNGTVTFGINTTSG-LKAG---KSTLNDGGLSIKNPTGSEQIQVGADGV 832

Qy        544 KFGRVKNWQVTDGEASSKLDFSKVIQRVAETEGTDEIGLIVNAKAGNDDIFVGQGKMNID 603
              || :| |     :|    : :| :  | |        ||||            | |     :
Db        833 KFAKVNN----NGVVGAGIDGTTRITR-------DEIG-----------FTG-----TN 864

Qy        604 GGDGHDRVFYSKDGGFGNITVDGTSAT--EAGSYTVNR--KVARGDIYHEVVKRQETKVG 659
              |     :    ||||     |     |    |   ::|    |     |   | || ::    | |:
Db        865 GSLDKSKPHLSKDG----INAGGKKITNIQSGEIAQNSHDAVTGGKIY-DLKTELENKIS 919

Qy        660 KRTETIQYRDYELR---------KVGYGYQSTDNLKSVEEV-------IGSQFN------ 697
                :| |   :|                |    |  |  |  |: : :        | :: |
Db        920 STAKTAQNSLHEFSVADEQGNNFTVSNPYSSYDTSKTSDVITFAGENGITTKVNKGVVRV 979

Qy        698 --DVFKG---SKFNDIFHSGEGDDLLDGGAGDDRLFG---------GKGNDRLSGDEGD 742
                 |  ||     |       ::|:|   ::|     |  : :  |            ||: |  : ::|:
Db        980 GIDQTKGLTTPKLTVGNNNGKG-IVIDSQNGQNTITGLSNTLANVTNDKGSVRTT-EQGN 1037

Qy        743 DLLD-----GGSGDDVLNGG---AGN---------DVYIFRKGDGNDT----LYDGTGN 780
               : |          |   |||: |     ||           |     |   |||  |        || |
Db       1038 IIKDEDKTRAASIVDVLSAGFNLQGNGEAVDFVSTYDTVNF--ADGNATTAKVTYDDTSK 1095

Qy        781 DKLAFADANISDIMIE--RTKEGI-----------------------IVKRNDHSGSI 813
                    | |:  |  ||    | |:                        :|| :|     :
Db       1096 TSKVVYDVNVDDTTIEVKDKKLGVKTTTLTSTGTGANKFALSNQATGDALVKASDIVAHL 1155

Qy        814 NIPRWYITSNLQNYQSNKTDHKIEQLIGKDGSYITSDQID-KILQDKKDGTVITSQELKK 872
              |     :: ::|      :           :    ||: :     |   |    |  ||||    ::|: |
Db       1156 NT----LSGDIQTAKGASQANNSAGYVDADGNKVIYDSTDNKYYQAKNDGTVDKTKEVAK 1211

Qy        873 --------------------------LADENKSQKLSASDI-------ASSLNKLVGSM 898
                                          : | || | :: :       |:| ||    :
Db       1212 DKLVAQAQTPDGTLAQMNVKSVINKEQVNDANKKQGINEDNAFVKGLEKAASDNKTKNAA 1271

Qy        899 ALFGTANSVSSNAL 912
                  |  |:|:   |
Db       1272 VTVGDLNAVAQTPL 1285
```

RESULT 20
US-10-175-275-3
; Sequence 3, Application US/10175275
; Publication No. US20030171254A1
; GENERAL INFORMATION:
; APPLICANT: SASAKI, Ken
; APPLICANT: HARKNESS, Robin E.
; APPLICANT: LOOSMORE, Sheena M.
; APPLICANT: CHONG, Pele
; APPLICANT: KLEIN, Michel H.
; TITLE OF INVENTION: HIGH MOLECULAR WEIGHT MAJOR OUTER MEMBRANE PROTEIN OF
; TITLE OF INVENTION: MORAXELLA
; FILE REFERENCE: 1038-1235 MIS
; CURRENT APPLICATION NUMBER: US/10/175,275
; CURRENT FILING DATE: 2002-06-20
; PRIOR APPLICATION NUMBER: 08/945,567
; PRIOR FILING DATE: 1998-03-19
; PRIOR APPLICATION NUMBER: 08/431,718
; PRIOR FILING DATE: 1995-05-01
; PRIOR APPLICATION NUMBER: 08/478,370
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 08/621,944
; PRIOR FILING DATE: 1996-03-26
; PRIOR APPLICATION NUMBER: PCT/CA96/00264
; PRIOR FILING DATE: 1996-04-29
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
;   LENGTH: 1992
;   TYPE: PRT
;   ORGANISM: Moraxella catarrhalis
US-10-175-275-3

```
  Query Match             4.2%;  Score 197;  DB 12;  Length 1992;
  Best Local Similarity   19.6%;  Pred. No. 8.3e-05;
  Matches  238;  Conservative  157;  Mismatches  407;  Indels  412;  Gaps   61;

Qy          2 SNINVIK----SNIQAGLNSTKSGLKNLYLAIPKDYDPQKGGTLNDFIKAADELGIARLA 57
               :|| |:|     | ::  |    |  :  |  :           | |: :     |: |  :
Db        340 NNIGVVKEADNSGLKVKLAKTLNNLTEVNTTTLNATTTVKVGSSSS--TTAELLSDSLTF 397

Qy         58 EEPN-HTETAKKSVDTVN--QFLSLTQTGIAISATKLEKFLQKHSTNKLAKGLD------ 108
               :||  :::  |:|  ||   :| :  :|   ||  ||: :         :|:   |
Db        398 TQPNTGSQSTSKTVYGVNGVKFTNNAETTAAIGTTRITR-------DKIGFARDGDVDEK 450

Qy        109 SVENIDRKLGKASNVLSTLSSFLGTALAGIELDSLIKKGDAAPDALAKASIDLINEIIGN 168
                  :|:|   |  :|  |: : :   || :   | : || :| ||:              |
Db        451 QAPYLDKKQLKVGSVAITIDNGID---AGNKKISNLAKGSSANDAVT----------IEQ 497

Qy        169 LSQSTQTIEAFSSQLAKLGSTISQAKGFSNIGNKLQNLNFSKTNLGLEIITGLLSGISAG 228
              |  :   |: |           |: ||          ::   |   |:       |:|::            | |
Db        498 LKAAKPTLNA--------GAGISVTPTEISVDAKSGNVTAPTYNIGVKTTELNSDGTSDK 549

Qy        229 FALADKNASTGKKVAAGFELS---NQVIGNVTKAISSYVL-------AQRVAAGLSTT-- 276
              |::   | :   |    :    |:      |:  |   | :              :       ||
Db        550 FSV--KGSGTNNSLVTAEHLASYLNEVNRTADSALQSFTVKEEDDDDANAITVAKDTTKN 607

Qy        277 -GAVAAL-------------------------------------------ITSSIMLAI 291
               |||: |                                             | :
Db        608 AGAVSILKLKGKNGLTVATKKDGTVTFGLSQDSGLTIGKSTLNNDGLTVKDTNEQIQVGA 667

Qy        292 SPLAFMNAADKFNHANALDEFAKQFR----KFGYDGDHLLAEYQRGVGTIEASLTTISTA 347
               : : ||     | :|      |    |      | |: |            | |  : |
Db        668 NGIKFTNV----NGSNPGTGIANTARITRDKIGFAGS------DGAVDTNKPYLDQDKLQ 717

Qy        348 LGAV---SAGVSAAAVGSAVGTPIALLVAGVTGLISGILEASKQAMFESVANRLQGKILE 404
              :| |    : |::|    ||             : |:: :    |  : | : :    : |: ||
Db        718 VGNVKITNTGINAG--GKA--------ITGLSPTLPSIADQSSRNI--ELGNTIQDK--- 762

Qy        405 WEKQNGG--QNYFDKGYDSRYAAYLANN---LKFLSE------LNKELEAERVIAITQQR 453
               :| |        :  : |::     | ||    : |:|            |         |  :
Db        763 -DKSNAASINDILNTGFN------LKNNNNPIDFVSTYDIVDFANGNATTATVTHDTANK 815

Qy        454 -----WDNNIGE----LAG---------ITKLGERIKSGKAYAD-----AFEDG----- 484
                   :| |: :     ||          |||  :    :|      :     : ||
Db        816 TSKVVYDVNVDDTTIHLTGTDDNKKLGVKTTKLNKTSANGNTATNFNVNSSDEDALVNAK 875

Qy        485 ----------------------------KKVEAGSN-------------------ITL 495
                                          |||:  :|                   :||
Db        876 DIAENLNTLAKEIHTTKGTADTALQTFTVKKVDENNNADDANAITVGQKNANNQVNTLTL 935

Qy        496 DAKTGIIDISNSNGKKTQALHFTSPLLTAGTESRERLTNGKYSYINK-----------L 543
                : |:   :: ||  |  :: ||  | ||   :  | :|  |  |             :
Db        936 KGENGLNIKTDKNGTVTFGINTTSG-LKAG---KSTLNDGGLSIKNPTGSEQIQVGADGV 991

Qy        544 KFGRVKNWQVTDGEASSKLDFSKVIQRVAETEGTDEIGLIVNAKAGNDDIFVGQGKMNID 603
              || :| |     :|    : :| :  | |        ||||            | |     :
Db        992 KFAKVNN----NGVVGAGIDGTTRITR-------DEIG-----------FTG-----TN 1023

Qy        604 GGDGHDRVFYSKDGGFGNITVDGTSAT--EAGSYTVNR--KVARGDIYHEVVKRQETKVG 659
              |     :    ||||     |     |    |   ::|    |     |   | || ::    | |:
Db       1024 GSLDKSKPHLSKDG----INAGGKKITNIQSGEIAQNSHDAVTGGKIY-DLKTELENKIS 1078

Qy        660 KRTETIQYRDYELR---------KVGYGYQSTDNLKSVEEV-------IGSQFN------ 697
                :| |   :|                |    |  |  |  |: : :        | :: |
Db       1079 STAKTAQNSLHEFSVADEQGNNFTVSNPYSSYDTSKTSDVITFAGENGITTKVNKGVVRV 1138
```

```
Qy        698 --DVFKG---SKFNDIFHSGEGDDLLDGGAGDDRLFG----------GKGNDRLSGDEGD 742
Db       1139 GIDQTKGLTTPKLTVGNNNGKG-IVIDSQNGQNTITGLSNTLANVTNDKGSVRTT-EQGN 1196

Qy        743 DLLD-----GGSGDDVLNGG---AGN----------DVYIFRKGDGNDT----LYDGTGN 780
Db       1197 IIKDEDKTRAASIVDVLSAGFNLQGNGEAVDFVSTYDTVNF--ADGNATTAKVTYDDTSK 1254

Qy        781 DKLAFADANISDIMIE--RTKEGI------------------------IVKRNDHSGSI 813
Db       1255 TSKVVYDVNVDDTTIEVKDKKLGVKTTTLTSTGTGANKFALSNQATGDALVKASDIVAHL 1314

Qy        814 NIPRWYITSNLQNYQSNKTDHKIEQLIGKDGSYITSDQID-KILQDKKDGTVITSQELKK 872
Db       1315 NT----LSGDIQTAKGASQANNSAGYVDADGNKVIYDSTDNKYYQAKNDGTVDKTKEVAK 1370

Qy        873 --------------------------LADENKSQKLSASDI-------ASSLNKLVGSM 898
Db       1371 DKLVAQAQTPDGTLAQMNVKSVINKEQVNDANKKQGINEDNAFVKGLEKAASDNKTKNAA 1430

Qy        899 ALFGTANSVSSNAL 912
Db       1431 VTVGDLNAVAQTPL 1444
```

RESULT 21
US-10-175-282-3
; Sequence 3, Application US/10175282
; Publication No. US20030170657A1
; GENERAL INFORMATION:
; APPLICANT: SASAKI, Ken
; APPLICANT: HARKNESS, Robin E.
; APPLICANT: LOOSMORE, Sheena M.
; APPLICANT: CHONG, Pele
; APPLICANT: KLEIN, Michel H.
; TITLE OF INVENTION: HIGH MOLECULAR WEIGHT MAJOR OUTER MEMBRANE PROTEIN OF
; TITLE OF INVENTION: MORAXELLA
; FILE REFERENCE: 1038-1234 MIS
; CURRENT APPLICATION NUMBER: US/10/175,282
; CURRENT FILING DATE: 2002-06-20
; PRIOR APPLICATION NUMBER: 08/945,567
; PRIOR FILING DATE: 1998-03-19
; PRIOR APPLICATION NUMBER: 08/431,718
; PRIOR FILING DATE: 1995-05-01
; PRIOR APPLICATION NUMBER: 08/478,370
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 08/621,944
; PRIOR FILING DATE: 1996-03-26
; PRIOR APPLICATION NUMBER: PCT/CA96/00264
; PRIOR FILING DATE: 1996-04-29
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
;   LENGTH: 1992
;   TYPE: PRT
;   ORGANISM: Moraxella catarrhalis
US-10-175-282-3

```
  Query Match             4.2%;  Score 197;  DB 12;  Length 1992;
  Best Local Similarity   19.6%;  Pred. No. 8.3e-05;
  Matches  238;  Conservative  157;  Mismatches  407;  Indels  412;  Gaps   61;

Qy          2 SNINVIK----SNIQAGLNSTKSGLKNLYLAIPKDYDPQKGGTLNDFIKAADELGIARLA 57
Db        340 NNIGVVKEADNSGLKVKLAKTLNNLTEVNTTTLNATTTVKVGSSSS--TTAELLSDSLTF 397

Qy         58 EEPN-HTETAKKSVDTVN--QFLSLTQTGIAISATKLEKFLQKHSTNKLAKGLD------ 108
Db        398 TQPNTGSQSTSKTVYGVNGVKFTNNAETTAAIGTTRITR-------DKIGFARDGDVDEK 450

Qy        109 SVENIDRKLGKASNVLSTLSSFLGTALAGIELDSLIKKGDAAPDALAKASIDLINEIIGN 168
Db        451 QAPYLDKKQLKVGSVAITIDNGID---AGNKKISNLAKGSSANDAVT----------IEQ 497

Qy        169 LSQSTQTIEAFSSQLAKLGSTISQAKGFSNIGNKLQNLNFSKTNLGLEIITGLLSGISAG 228
Db        498 LKAAKPTLNA--------GAGISVTPTEISVDAKSGNVTAPTYNIGVKTTELNSDGTSDK 549

Qy        229 FALADKNASTGKKVAAGFELS---NQVIGNVTKAISSYVL-------AQRVAAGLSTT-- 276
Db        550 FSV--KGSGTNNSLVTAEHLASYLNEVNRTADSALQSFTVKEEDDDDANAITVAKDTTKN 607

Qy        277 -GAVAAL--------------------------------------------ITSSIMLAI 291
Db        608 AGAVSILKLKGKNGLTVATKKDGTVTFGLSQDSGLTIGKSTLNNDGLTVKDTNEQIQVGA 667

Qy        292 SPLAFMNAADKFNHANALDEFAKQFR----KFGYDGDHLLAEYQRGVGTIEASLTTISTA 347
Db        668 NGIKFTNV----NGSNPGTGIANTARITRDKIGFAGS------DGAVDTNKPYLDQDKLQ 717

Qy        348 LGAV---SAGVSAAAVGSAVGTPIALLVAGVTGLISGILEASKQAMFESVANRLQGKILE 404
Db        718 VGNVKITNTGINAG--GKA---------ITGLSPTLPSIADQSSRNI--ELGNTIQDK--- 762

Qy        405 WEKQNGG--QNYFDKGYDSRYAAYLANN---LKFLSE------LNKELEAERVIAITQQR 453
Db        763 -DKSNAASINDILNTGFN------LKNNNNPIDFVSTYDIVDFANGNATTATVTHDTANK 815

Qy        454 -----WDNNIGE----LAG----------ITKLGERIKSGKAYAD-----AFEDG----- 484
Db        816 TSKVVYDVNVDDTTIHLTGTDDNKKLGVKTTKLNKTSANGNTATNFNVNSSDEDALVNAK 875

Qy        485 -----------------------------KKVEAGSN-------------------ITL 495
Db        876 DIAENLNTLAKEIHTTKGTADTALQTFTVKKVDENNNADDANAITVGQKNANNQVNTLTL 935

Qy        496 DAKTGIIDISNSNGKKTQALHFTSPLLTAGTESRERLTNGKYSYINK-----------L 543
Db        936 KGENGLNIKTDKNGTVTFGINTTSG-LKAG---KSTLNDGGLSIKNPTGSEQIQVGADGV 991

Qy        544 KFGRVKNWQVTDGEASSKLDFSKVIQRVAETEGTDEIGLIVNAKAGNDDIFVGQGKMNID 603
Db        992 KFAKVNN----NGVVGAGIDGTTRITR-------DEIG------------FTG-----TN 1023

Qy        604 GGDGHDRVFYSKDGGFGNITVDGTSAT--EAGSYTVNR--KVARGDIYHEVVKRQETKVG 659
Db       1024 GSLDKSKPHLSKDG----INAGGKKITNIQSGEIAQNSHDAVTGGKIY-DLKTELENKIS 1078

Qy        660 KRTETIQYRDYELR---------KVGYGYQSTDNLKSVEEV-------IGSQFN------ 697
Db       1079 STAKTAQNSLHEFSVADEQGNNFTVSNPYSSYDTSKTSDVITFAGENGITTKVNKGVVRV 1138

Qy        698 --DVFKG---SKFNDIFHSGEGDDLLDGGAGDDRLFG----------GKGNDRLSGDEGD 742
Db       1139 GIDQTKGLTTPKLTVGNNNGKG-IVIDSQNGQNTITGLSNTLANVTNDKGSVRTT-EQGN 1196

Qy        743 DLLD-----GGSGDDVLNGG---AGN----------DVYIFRKGDGNDT----LYDGTGN 780
Db       1197 IIKDEDKTRAASIVDVLSAGFNLQGNGEAVDFVSTYDTVNF--ADGNATTAKVTYDDTSK 1254

Qy        781 DKLAFADANISDIMIE--RTKEGI------------------------IVKRNDHSGSI 813
Db       1255 TSKVVYDVNVDDTTIEVKDKKLGVKTTTLTSTGTGANKFALSNQATGDALVKASDIVAHL 1314

Qy        814 NIPRWYITSNLQNYQSNKTDHKIEQLIGKDGSYITSDQID-KILQDKKDGTVITSQELKK 872
Db       1315 NT----LSGDIQTAKGASQANNSAGYVDADGNKVIYDSTDNKYYQAKNDGTVDKTKEVAK 1370

Qy        873 --------------------------LADENKSQKLSASDI-------ASSLNKLVGSM 898
Db       1371 DKLVAQAQTPDGTLAQMNVKSVINKEQVNDANKKQGINEDNAFVKGLEKAASDNKTKNAA 1430

Qy        899 ALFGTANSVSSNAL 912
Db       1431 VTVGDLNAVAQTPL 1444
```

```
RESULT 22
US-09-815-242-5639
; Sequence 5639, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
;  APPLICANT: Haselbeck, Robert
;  APPLICANT:  Ohlsen, Kari L.
;  APPLICANT:  Zyskind, Judith W.
;  APPLICANT:  Wall, Daniel
;  APPLICANT:  Trawick, John D.
;  APPLICANT:  Carr, Grant J.
;  APPLICANT:  Yamamoto, Robert T.
;  APPLICANT:  Xu, H. Howard
;  TITLE OF INVENTION: Identification of Essential Genes in
;  TITLE OF INVENTION:  Prokaryotes
;  FILE REFERENCE: ELITRA.011A
;  CURRENT APPLICATION NUMBER: US/09/815,242
;  CURRENT FILING DATE:  2001-03-21
;  PRIOR APPLICATION NUMBER: 60/191,078
;  PRIOR FILING DATE: 2000-03-21
;  PRIOR APPLICATION NUMBER: 60/206,848
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 60/207,727
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: 60/242,578
;  PRIOR FILING DATE: 2000-10-23
;  PRIOR APPLICATION NUMBER: 60/253,625
;  PRIOR FILING DATE: 2000-11-27
;  PRIOR APPLICATION NUMBER: 60/257,931
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: 60/269,308
;  PRIOR FILING DATE: 2001-02-16
;  NUMBER OF SEQ ID NOS: 14110
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5639
;   LENGTH: 2086
;   TYPE: PRT
;   ORGANISM: Staphylococcus aureus
US-09-815-242-5639

  Query Match              4.2%;  Score 196;  DB 9;  Length 2086;
  Best Local Similarity   18.6%;  Pred. No. 0.0001;
  Matches  221;  Conservative  165;  Mismatches  431;  Indels  372;  Gaps   54;

Qy          1 MSNINVIKSNIQAGLNSTKSGLKNLYLAIPKDYDPQKGGTLNDFIKAADELGIARLAEEP 60
Db        660 LEDIQTVQTNGQA-LNNAMKGLRD---SIANE----------TTVKASQNYTDA----SP 701

Qy         61 NHTETAKKSVDT----VNQFLSLTQTGIAISATKLEKFLQKHSTNKLAKGLDSVENIDRK 116
Db        702 NNQSTYNSAVSNAKGIINQTNNPTMDTSAITQATTQ-------VNNAKNGLNGAENLRNA 754

Qy        117 LGKASNVLSTLSSFLGTALAGIELDSLIKKGDAAPDALAKASIDLINEIIGNLSQS---- 172
Db        755 QNTAKQNLNTLSHLTNNQKSAIS-SQIDRAGHVSEVTAAKNAATELNTQMGNLEQAIHDQ 813

Qy        173 ---TQTIEAFSSQLAKLGSTISQAKGFSNIGNKLQNLNFSKTNL--GLEIITGLLSGISA 227
Db        814 NTVKQGVNFTDADKAKRDAYTNAVSRAETILNKTQGANTSKQDVEAAIQNVTSAKNALN- 872

Qy        228 GFALADKNASTGKKVAAGFELSNQVIGNVT-------KAISSYVLAQRVAAG---LSTTG 277
Db        873 ----GDQNVTNAKNTA------KHALNNLTSINNAQKRDLTTKIDQATTVAGVEAVSNTG 922

Qy        278 AVAALITSSIMLAISPLAFMNAADKFNHANALDEFAKQFRKFGYDGDHLLAEYQRGVGT- 336
Db        923 TQLNTAMANLQNGINDKANTLASENYHDADSDKKTA--YTQAVTNAENIL---NKNSGSN 977

Qy        337 -----IEASLTTISTALGAVSAGVSAAAVGSAVGTPIALLVAGVTGLISGILEASKQAM- 390
Db        978 LDKAAVENALSQVTNAKGALNGNHNLEQAKSNANTTINGLQHLTTAQKDKLKQQVQQAQN 1037

Qy        391 ------FESVANRLQG-------KILEWEKQNGGQNYFD----------KGYDSRYAAYL 427
Db       1038 VAGVDTVKSSANTLNGAMGTLRNSIQDNTATNNGQNYLDATESNKTNYNNAVDSANGVIN 1097

Qy        428 A-------------------------------------NNLKFLSELNKE-------- 440
Db       1098 ATSNPNMDANAINQIATQVTSTKNALDGTHNLTQAKQTATNAIDGATNLNKAQKDALKAQ 1157

Qy        441 -LEAERVIAITQ-QRWDNNIGELAGITKLG----ERIKSGKAYADAFEDGKKV------- 487
Db       1158 VTSAQRVANVTSIQQTANELNTAMGQLQHGIDDENATKQTQKYRDA-EQSKKTAYDQAVA 1216

Qy        488 --------EAGSN-------------ITLDAKTGIIDISNSNGKKTQAL----HFTSPL 521
Db       1217 AAKAILNKQTGSNSDKAAVDRALQQVTSTKDALNGDAKLAEAKAAAKQNLGTLNHITNAQ 1276

Qy        522 LTA-----------GTESRERLTNGKYSYINKL------KFGRVKNWQVTDGEASSKLD 563
Db       1277 RTALEGQINQATTVDGVNTVKTNANTLDGAMNSLQGSINDKDATLRNQNYLDADESKRNA 1336

Qy        564 FSKVIQRVAETEGTDEIGLIVNAKAGNDDIFVGQGKMNIDGGDGHDRVFYSKD--GGFGN 621
Db       1337 YT---QAVTAAEG------ILNKQTGGN----TSKADVD--NALNTVTRAKAALNGAEN 1380

Qy        622 ITVDGTSATEAGSYTVN-----RKVARGDIYHEVVKRQ--------------ETKVGKR 661
Db       1381 LRNTKTSATN----TINGLPNLTQLQKDNLKHQVEQAQNVAGVNGVKDKGNTLNTAMGAL 1436

Qy        662 TETIQ----------YRD-YELRKVGY--------------------------------- 677
Db       1437 RTSIQNDNTTKTSQNYLDASDSNKNNYNTAVNNANGVINVTNNPNMDANAINGMANQVNT 1496

Qy        678 ------GYQ--------STDNLKSVEEVIGSQFN-----------------DVFKGSKFN 706
Db       1497 TKAALNGAQNLAQAKTTAKNNLNNLTSINNAQKDALTRSIDGATTVAGVNQETAKATELN 1556

Qy        707 DIFHSGEGDDLLDGGAGDDRLFGGKGNDRLSGDEGDDLLDGGSG-----DDVLNGGAGND 761
Db       1557 NAMHS------LQNGI----------NDETQTKQTQKYLDAGPSKKSAYDQAVNAAKA-- 1598

Qy        762 VYIFRKGDGNDTLYDGTGNDKLAFADANISDIMIERTKEGIIVKRNDHSGSINIPRWYIT 821
Db       1599 --ILTKASGQNV-------DKAAVEQA-LQNVNSTKTALNGDAKLNEAKAAAKQTLGTLT 1648

Qy        822 SNLQNYQSNKTDHKIEQLIGKDGSYITS----------DQIDKILQDKKDGTVITSQELK 871
Db       1649 -HINNAQRTALDNEITQATNVEGVNTVKAKAQQLDGAMGQLETSIRDK--DTTLQSQNYQ 1705

Qy        872 KLADENK---SQKLSASDIASSLNKLVGSMALFGTANSVSSNALQPITQ 917
Db       1706 DADDAKRTAYSQAVNAA--ATILNKTAGG----NTPKADVERAMQAVTQ 1748

RESULT 23
US-09-815-242-12610
; Sequence 12610, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
;  APPLICANT: Haselbeck, Robert
;  APPLICANT:  Ohlsen, Kari L.
;  APPLICANT:  Zyskind, Judith W.
;  APPLICANT:  Wall, Daniel
;  APPLICANT:  Trawick, John D.
;  APPLICANT:  Carr, Grant J.
;  APPLICANT:  Yamamoto, Robert T.
;  APPLICANT:  Xu, H. Howard
;  TITLE OF INVENTION: Identification of Essential Genes in
;  TITLE OF INVENTION:  Prokaryotes
;  FILE REFERENCE: ELITRA.011A
;  CURRENT APPLICATION NUMBER: US/09/815,242
;  CURRENT FILING DATE:  2001-03-21
;  PRIOR APPLICATION NUMBER: 60/191,078
```

```
;  PRIOR FILING DATE: 2000-03-21
;  PRIOR APPLICATION NUMBER: 60/206,848
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 60/207,727
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: 60/242,578
;  PRIOR FILING DATE: 2000-10-23
;  PRIOR APPLICATION NUMBER: 60/253,625
;  PRIOR FILING DATE: 2000-11-27
;  PRIOR APPLICATION NUMBER: 60/257,931
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: 60/269,308
;  PRIOR FILING DATE: 2001-02-16
;  NUMBER OF SEQ ID NOS: 14110
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 12610
;   LENGTH: 5795
;   TYPE: PRT
;   ORGANISM: Staphylococcus aureus
US-09-815-242-12610

  Query Match             4.2%;  Score 196;  DB 9;  Length 5795;
  Best Local Similarity   18.6%;  Pred. No. 0.0005;
  Matches  221;  Conservative  165;  Mismatches  431;  Indels  372;  Gaps   54;

Qy          1 MSNINVIKSNIQAGLNSTKSGLKNLYLAIPKDYDPQKGGTLNDFIKAADELGIARLAEEP 60
Db       2937 LEDIQTVQTNGQA-LNNAMKGLRD---SIANE----------TTVKASQNYTDA----SP 2978

Qy         61 NHTETAKKSVDT----VNQFLSLTQTGIAISATKLEKFLQKHSTNKLAKGLDSVENIDRK 116
Db       2979 NNQSTYNSAVSNAKGIINQTNNPTMDTSAITQATTQ-------VNNAKNGLNGAENLRNA 3031

Qy        117 LGKASNVLSTLSSFLGTALAGIELDSLIKKGDAAPDALAKASIDLINEIIGNLSQS---- 172
Db       3032 QNTAKQNLNTLSHLTNNQKSAIS-SQIDRAGHVSEVTAAKNAATELNTQMGNLEQAIHDQ 3090

Qy        173 ---TQTIEAFSSQLAKLGSTISQAKGFSNIGNKLQNLNFSKTNL--GLEIITGLLSGISA 227
Db       3091 NTVKQGVNFTDADKAKRDAYTNAVSRAETILNKTQGANTSKQDVEAAIQNVTSAKNALN- 3149

Qy        228 GFALADKNASTGKKVAAGFELSNQVIGNVT-------KAISSYVLAQRVAAG---LSTTG 277
Db       3150 ----GDQNVTNAKNTA------KHALNNLTSINNAQKRDLTTKIDQATTVAGVEAVSNTG 3199

Qy        278 AVAALITSSIMLAISPLAFMNAADKFNHANALDEFAKQFRKFGYDGDHLLAEYQRGVGT- 336
Db       3200 TQLNTAMANLQNGINDKANTLASENYHDADSDKKTA--YTQAVTNAENIL---NKNSGSN 3254

Qy        337 -----IEASLTTISTALGAVSAGVSAAAVGSAVGTPIALLVAGVTGLISGILEASKQAM- 390
Db       3255 LDKAAVENALSQVTNAKGALNGNHNLEQAKSNANTTINGLQHLTTAQKDKLKQQVQQAQN 3314

Qy        391 ------FESVANRLQG-------KILEWEKQNGGQNYFD----------KGYDSRYAAYL 427
Db       3315 VAGVDTVKSSANTLNGAMGTLRNSIQDNTATNNGQNYLDATESNKTNYNNAVDSANGVIN 3374

Qy        428 A--------------------------------------NNLKFLSELNKE-------- 440
Db       3375 ATSNPNMDANAINQIATQVTSTKNALDGTHNLTQAKQTATNAIDGATNLNKAQKDALKAQ 3434

Qy        441 -LEAERVIAITQ-QRWDNNIGELAGITKLG----ERIKSGKAYADAFEDGKKV------- 487
Db       3435 VTSAQRVANVTSIQQTANELNTAMGQLQHGIDDENATKQTQKYRDA-EQSKKTAYDQAVA 3493

Qy        488 --------EAGSN--------------ITLDAKTGIIDISNSNGKKTQAL----HFTSPL 521
Db       3494 AAKAILNKQTGSNSDKAAVDRALQQVTSTKDALNGDAKLAEAKAAAKQNLGTLNHITNAQ 3553

Qy        522 LTA------------GTESRERLTNGKYSYINKL------KFGRVKNWQVTDGEASSKLD 563
Db       3554 RTALEGQINQATTVDGVNTVKTNANTLDGAMNSLQGSINDKDATLRNQNYLDADESKRNA 3613

Qy        564 FSKVIQRVAETEGTDEIGLIVNAKAGNDDIFVGQGKMNIDGGDGHDRVFYSKD--GGFGN 621
Db       3614 YT---QAVTAAEG------ILNKQTGGN-----TSKADVD--NALNTVTRAKAALNGAEN 3657

Qy        622 ITVDGTSATEAGSYTVN-----RKVARGDIYHEVVKRQ---------------ETKVGKR 661
Db       3658 LRNTKTSATN----TINGLPNLTQLQKDNLKHQVEQAQNVAGVNGVKDKGNTLNTAMGAL 3713

Qy        662 TETIQ---------YRD-YELRKVGY---------------------------------- 677
Db       3714 RTSIQNDNTTKTSQNYLDASDSNKNNYNTAVNNANGVINVTNNPNMDANAINGMANQVNT 3773

Qy        678 ------GYQ--------STDNLKSVEEVIGSQFN-----------------DVFKGSKFN 706
Db       3774 TKAALNGAQNLAQAKTTAKNNLNNLTSINNAQKDALTRSIDGATTVAGVNQETAKATELN 3833

Qy        707 DIFHSGEGDDLLDGGAGDDRLFGGKGNDRLSGDEGDDLLDGGSG-----DDVLNGGAGND 761
Db       3834 NAMHS------LQNGI----------NDETQTKQTQKYLDAGPSKKSAYDQAVNAAKA-- 3875

Qy        762 VYIFRKGDGNDTLYDGTGNDKLAFADANISDIMIERTKEGIIVKRNDHSGSINIPRWYIT 821
Db       3876 --ILTKASGQNV-------DKAAVEQA-LQNVNSTKTALNGDAKLNEAKAAAKQTLGTLT 3925

Qy        822 SNLQNYQSNKTDHKIEQLIGKDGSYITS----------DQIDKILQDKKDGTVITSQELK 871
Db       3926 -HINNAQRTALDNEITQATNVEGVNTVKAKAQQLDGAMGQLETSIRDK--DTTLQSQNYQ 3982

Qy        872 KLADENK---SQKLSASDIASSLNKLVGSMALFGTANSVSSNALQPITQ 917
Db       3983 DADDAKRTAYSQAVNAA--ATILNKTAGG----NTPKADVERAMQAVTQ 4025

RESULT 24
US-10-369-493-19792
; Sequence 19792, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
;  APPLICANT: Cao, Yongwei
;  APPLICANT:  Hinkle, Gregory J.
;  APPLICANT:  Slater, Steven C.
;  APPLICANT:  Goldman, Barry S.
;  APPLICANT:  Chen, Xianfeng
;  TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
;  TITLE OF INVENTION:  PLANTS WITH IMPROVED PROPERTIES
;  FILE REFERENCE: 38-10(52052)B
;  CURRENT APPLICATION NUMBER: US/10/369,493
;  CURRENT FILING DATE:  2003-02-28
;  PRIOR APPLICATION NUMBER: US 60/360,039
;  PRIOR FILING DATE: 2002-02-21
;  NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 19792
;   LENGTH: 262
;   TYPE: PRT
;   ORGANISM: No. US20030233675A1toc punctiforme
US-10-369-493-19792

  Query Match             4.1%;  Score 191;  DB 12;  Length 262;
  Best Local Similarity   26.2%;  Pred. No. 1e-05;
  Matches   70;  Conservative   35;  Mismatches   80;  Indels   82;  Gaps   13;

Qy        532 LTNGKYSYINKLKFGRVKNWQVTDGE---ASSKLDFSKVIQRVAETEGTDEIGLIVNAKA 588
Db         16 LTGVKPTFLGSFGSSNSDNITVQPGKTFFAGDGADF------VEGTEGN-------TIQA 62

Qy        589 GNDD--IFVGQGKMNIDGGDGHDRVFYSKDGGFGNITVDGTSATEAGSYTVNRKVARGDI 646
Db         63 GNDDDTVLVGSNS-SVSAGDGNDQIVIGQNGAAQNISADG---------------GNGDD 106

Qy        647 YHEVVKRQETKVGKRTETIQYRDYELRKVGYGYQSTDNLKSVEEVIGSQFND---VFKGS 703
Db        107 VITVVEA-------------------------NGNNNL------FGAEGNDTLTVIEGS 134
```

```
Qy        704 KFNDIFHSGEGDDLLDGGAGDDRLFGGKGNDRLSGDEGDDLLDGGSGDDVLNGGAGNDVY 763
Db        135 R-QSLF-GGLGNDTLTSNGSNNRLYGGSGDDKIFSNTNDSLF-GGDGDDVLFAGQG---- 187

Qy        764 IFRKGDGNDTLYDGTGNDKLAFADANI 790
Db        188 ------GSNRLSGGAGADQFWIANASL 208
```

RESULT 25
US-09-797-862-32
; Sequence 32, Application US/09797862
; Patent No. US20020102276A1
; GENERAL INFORMATION:
; APPLICANT: PEAK, IAN RICHARD ANSELM
; APPLICANT: JENNINGS, MICHAEL PAUL
; APPLICANT: MOXON, E. RICHARD
; TITLE OF INVENTION: NOVEL SURFACE ANTIGEN
; FILE REFERENCE: 065064/0134
; CURRENT APPLICATION NUMBER: US/09/797,862
; CURRENT FILING DATE: 2001-05-03
; PRIOR APPLICATION NUMBER: PCT/AU98/01031
; PRIOR FILING DATE: 1998-12-14
; PRIOR APPLICATION NUMBER: GB 9726398.2
; PRIOR FILING DATE: 1997-12-12
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 32
; LENGTH: 1098
; TYPE: PRT
; ORGANISM: Haemophilus influenzae
US-09-797-862-32

Query Match 4.1%; Score 188.5; DB 10; Length 1098;
Best Local Similarity 19.7%; Pred. No. 0.00014;
Matches 225; Conservative 156; Mismatches 389; Indels 373; Gaps 59;

```
Qy          5 NVIKSNIQAGLNSTKSGLKNLYLAIPKDYDPQKGGTLNDFIKAADELGIARLAEEPNHTE 64
Db         57 NKLKAYGDANFNFTNNSIADAEKQVQEAYK----GLLNLNEKNASD----KLLVEDNTAA 108

Qy         65 TAKKSVDTVNQFLSLTQTGIAISA---TKLEKFLQ-KHSTNKLAKGLDSVENIDRKLGKA 120
Db        109 TVG---------NLRKLGWVLSSKNGTRNEKSQQVKHADEVLFEGKGGVQVTSTSENGK 158

Qy        121 SNVLSTLSSFLGTALAGIELDSLIKKGDAAPDALAKASIDLINEIIGNLSQSTQTIEAFS 180
Db        159 HTITFALAKDLGVKTATVS-DTLTIGGGAAAGATTTPKV--------NVTSTTDGLKFAK 209

Qy        181 SQLAKLGSTISQAKGFSNIGNKL----------------QNLNFSKT-------NLGLEI 217
Db        210 DAAGANGDTTVHLNG---IGSTLTDTLVGSPATHIDGGDQSTHYTRAASIKDVLNAGWNI 266

Qy        218 ITGLLSGISAGFALADKNASTGKKVAAGFELSNQVIGNVTKAISSYVLAQRVAAGLSTT- 276
Db        267 -----KGVKAG-------STTGQSENVDF------------VHTYDTVEFLSADTETTT 301

Qy        277 ----------------GAVAALITSSIMLAISPLAFMNAADKFNHANALDEFAKQFRKFG 320
Db        302 VTVDSKENGKRTEVKIGAKTSVIKEK-----DGKLFTGKANKETNK-------------- 342

Qy        321 YDGDHLLAEYQRGVGTIEASLTTISTALGAVSAGVSAAAVGSAVGT--PIALLVAGVTGL 378
Db        343 VDGANATEDADEGKG-----LVTAKDVIDAVNKTGWRIKTTDANGQNGDFATVASG---- 393

Qy        379 ISGILEASKQAMFESVANRLQGKILEWEKQ--------------------NGGQNYFDK 417
Db        394 -TNVTFASGNGTTATVTNGTDGITVKYDAKVGDGLKLDGDKIAADTTALTVNDGKN---- 448

Qy        418 GYDSRYAAYLANNLK-----FLSELNKELEAER--VIAITQQRWDNNIGELAGITKLG-- 468
Db        449 ----------ANNPKGKVADVASTDEKKLVTAKGLVTALNSLSWTTTAAEADGGTLDGNA 498

Qy        469 --ERIKSGKAYADAFEDGKKV---EAGSNITL---DAKTGIIDI---SNSNGKKTQ---- 513
Db        499 SEQEVKAGDKV--TFKAGKNLKVKQEGANFTYSLQDALTGLTSITLGTGNNGAKTEINKD 556

Qy        514 ---------------ALHFTSPLLTAGTESRERLTNGKYSYINKLKFGRVKNWQVT--- 554
Db        557 GLTITPANGAGANNANTISVTKDGISAGGQSVKNVVSG------LKKFGDANFDPLTSSA 610

Qy        555 DGEASSKLDFSKVIQRVAETEGTDEIGLIV--NAKAGNDDI-----FVGQGKMNIDGGDG 607
Db        611 DNLTKQNDDAYKGLTNLDE-KGTDKQTPVVADNTAATVGDLRGLGWVISADKTTGGSTEY 669

Qy        608 HDRVFYSKDGGF--GN-ITVDGTSATEAGSYTVNRKVARGDIYHEVVKRQETKV----GK 660
Db        670 HDQVRNANEVKFKSGNGINVSG--KTVNGRREITFELAKG----EVVKSNEFTVKETNGK 723

Qy        661 RTETIQYRDYELRKVGYGYQSTDNLKSVE--------EVIGSQFNDVFKGSKFNDIFHSG 712
Db        724 ETSLV--------KVGDKYYSKEDIDLTTGQPKLKDGNTVAAKYQD--KGGKVVSVTDNT 773

Qy        713 EGDDLLDGGAGDDRLFGGKGNDRLSGDEGDDLLDGGSGDDVLNGGAGNDVYIFRK-GDGN 771
Db        774 EA-----------TITNKGSGYVTGNQVADAI-AKSGFEL---GLADEADAKRAFDDKT 817

Qy        772 DTLYDGT-----GNDKLAFA---DANISDIMIERTKEGIIVKRNDHSGSINIPRWYITSN 823
Db        818 KALSAGTTEIVNAHDKVRFANGLNTKVSAATVEST---------DANGD-KVTTTFVKTD 867

Qy        824 L-----QNYQSNKTDHKIEQLIGKDG----------------SYITSDQID----KILQD 858
Db        868 VELPLTQIYNTDANGKKITKVV-KDGQTKWYELNADGTADMTKEVTLGNVDSDGKKVVKD 926

Qy        859 --------KKDGT------------VITSQELKKLADENKSQK-------------LSA- 884
Db        927 NDGKWYHAKADGTADKTKGEVSNDKVSTDEKHVVSLDPNDQSKGKGVVIDNVANGDISAT 986

Qy        885 -----------------SDIASSLNKL------VGSMALFGTANSVSSNALQPITQPTQG 921
Db        987 STDAINGSQLYAVAKGVTNLAGQVNNLEGKVNKVGKRADAGTASALAASQLPQATMPGKS 1046

Qy        922 ILA 924
Db       1047 MVA 1049
```

RESULT 26
US-10-185-990-11
; Sequence 11, Application US/10185990
; Publication No. US20030073109A1
; GENERAL INFORMATION:
; APPLICANT: Pan, Jae-Gu
; TITLE OF INVENTION: JAE GU PAN ET AL
; FILE REFERENCE: 02589.000100
; CURRENT APPLICATION NUMBER: US/10/185,990
; CURRENT FILING DATE: 2002-06-28
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 11
; LENGTH: 1626
; TYPE: PRT
; ORGANISM: Bacillus subtilis
US-10-185-990-11

Query Match 4.0%; Score 188; DB 15; Length 1626;
Best Local Similarity 20.2%; Pred. No. 0.00027;
Matches 185; Conservative 112; Mismatches 333; Indels 286; Gaps 39;

```
Qy         63 TETAKKSVDTVNQFLSLTQTGIAISATKLEKFLQKHSTNKLAKGLDSVENIDRKLGKASN 122
Db        785 TQTAQERSDLVTGYGSTSTAGYASS---------------LIAGYGSTQ--------TAG 821

Qy        123 VLSTLSSFLGTALAGIELDSLIKKGDAAPDALAKASIDLI-----NEIIGNLSQ------ 171
```

```
Db        822 YESTLTAGYGSTQTAQE-NSSLTTGYGSTSTAGFAS-SLIAGYGSTQTAGYKSTLTAGYG 879
Qy        172 STQTIEAFSSQLAKLGSTISQAKGFSNI---GNKLQN--LNFSKTNLGLEIITGLLSGIS 226
Db        880 STQTAEYGSSLTAGYGSTATAGQDSSLIAGYGSSLTSGIRSFLTAGYGSTLIAGLRSVLI 939
Qy        227 AGFAL---------------ADKNASTGKKVAAGFELSNQVIGNVTKAISSYVLAQRVAA 271
Db        940 AGYGSSLTSGIRSTLTAGYGSNQIASYGSSLIAGHE-SIQVAGNKSMLIAGKGSSQ--TA 996
Qy        272 GLSTTGAVAALITSSIMLAISPLAFMNAADKFNHANALDEFAKQFRKFGYDGDHLLAEYQ 331
Db        997 GFRSTLIAGA---GSVQLAGDRSRLIAGADS-------NQTAGDRSKLLAGNNSYLTAGD 1046
Qy        332 RGVGTIEASLTTISTALGAVSAG----VSAAAVGSAVGTPIALLVAGV-TGLISGILEAS 386
Db       1047 RSKLTGGHDCTLMAGDQSRLTAGKNSVLTAGARSKLIGSEGSTLSAGEDSTLIFRLWDGK 1106
Qy        387 --KQAMFESVANRLQGKIL-----------------EWEKQNGGQNYFDKGYDSRYAAYL 427
Db       1107 RYRQLVARTGENGVEADIPYYVNEDDDIVDKPDEDDDWIEVKPGMGVFD---------- 1155
Qy        428 ANNLKFLSELNKEL--EAERVIAITQQRWDNN------------------IGELAGITKL 467
Db       1156 YKNLG--TEASKTLFADATAITLYTYHNLDNGFAVGYQQHGLGLGLPATLVGALLGST-- 1211
Qy        468 GERIKSGKAYADAFEDGKKVEAGSNITLDAKTGIIDISNSNG---KKTQALHFTSPLLTA 524
Db       1212 ---------------DSQGVIPGIPWNPDSEKAALDAVHAAGWTPISASALGYGGKVDAR 1256
Qy        525 GTESRERLTNGKYSYINKLKFGRVKNWQVTDGEASSKLDFSKVIQRVA----------ET 574
Db       1257 GT-----------------FFGEKAGYTTAQAEVLGKYDDAGKLLEIGIGFRGTSGPRES 1299
Qy        575 EGTDEIGLIVN---AKAGNDDI---FVGQ------------GKMNIDGGD----GHDRV 611
Db       1300 LITDSIGDLVSDLLAALGPKDYAKNYAGEAFGGLLKTVADYAGAHGLSGKDVLVSGHSL- 1358
Qy        612 FYSKDGGFG-NITVDGTSATEAGSYTVNRKVARGDIYHEVVKRQETKVGKRTETIQYRD- 669
Db       1359 -----GGLAVNSMADLSTSKWAGFYK--------DANYLAYASPTQSAGDKVLNIGYEND 1405
Qy        670 ----------YELRKVGY---GYQS-TDNLKSVEEVIGSQFNDVFKGSKFNDI------- 708
Db       1406 PVFRALDGSTFNLSSLGVHDKAHESTTDNIVS--------FNDHYASTLWNVLPFSIANL 1457
Qy        709 ------FHSGEGDDL---LDGG-----------------------------------AG 723
Db       1458 STWVSHLPSAYGDGMTRVLESGFYEQMTRDSTIIVANLSDPARANTWVQDLNRNAEPHTG 1517
Qy        724 DDRLFGGKGNDRLSGDEGDDLLDGGSGDDVLNGGAGNDVYIFRKGDGNDTLYDGTGNDKL 783
Db       1518 NTFIIGSDGNDLIQGGKGADFIEGGKGNDTIRDNSGHNTFLFSGHFGQDRIIGYQPTDRL 1577
Qy        784 AFADANISDIMIERTK 799
Db       1578 VFQGADGSTDLRDHAK 1593
```

RESULT 27
US-10-311-879-28
; Sequence 28, Application US/10311879
; Publication No. US20030186275A1
; GENERAL INFORMATION:
; APPLICANT: University of Sheffield
; TITLE OF INVENTION: Antigenic Peptides
; FILE REFERENCE: toxin
; CURRENT APPLICATION NUMBER: US/10/311,879
; CURRENT FILING DATE: 2003-03-17
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 28
;   LENGTH: 2659
;   TYPE: PRT
;   ORGANISM: Staphylococcus aureus
US-10-311-879-28

```
  Query Match             4.0%;  Score 186.5;  DB 12;  Length 2659;
  Best Local Similarity   19.0%;  Pred. No. 0.00073;
  Matches  186;  Conservative  129;  Mismatches  353;  Indels  309;  Gaps   44;

Qy          4 INVIKSNIQAGLNSTKSGLKNL-----YLAIPKDY---DPQKGGTLNDFIKAADELGIAR 55
Db       1784 VNTIKQN-ATNLNSAMGNLRQAVADKDQVKRTEDYADADTAKQNAYNSAVSSAETI---- 1838
Qy         56 LAEEPNHTETAKKSVDTVNQFLSLTQTGIAISATK-----LEKFLQKHSTNKLAKGLDSV 110
Db       1839 ----INQTTNPTMSVDDVNRATS------AVTSNKNALNGYEKLAQ--SKTDAARAIDAL 1886
Qy        111 ENIDR--------KLGKASNV-----LSTLSSFLGTALAGIE---------------LDS 142
Db       1887 PHLNNAQKADVKSKINAASNIAGVNTVKQQGTDLNTAMGNLQGAINDEQTTLNSQNYQDA 1946
Qy        143 LIKKGDAAPDALAKASIDLINEIIGNLSQSTQTIEAF---SSQLAKLGST--ISQAK--G 195
Db       1947 TPSKKTAYTNAV-QAAKDILNKSNGQNKTKDQVTEAMNQVNSAKNNLDGTRLLDQAKQTA 2005
Qy        196 FSNIGNKLQNLNFSKTNLGLEIITGLLSGISAGFALADKNASTGKKVAAGFELSNQVIGN 255
Db       2006 KQQLNNMTHLTTAQKTNLTNQINSGT---TVAGVQTVQSNANT---LDQAMNTLRQSIAN 2059
Qy        256 --VTKAISSYVLA--QRVAAGLSTTGAVAALITSSIMLAISPLAFMNAADKFNHANALDE 311
Db       2060 KDATKASEDYVDANNDKQTAYNNAVAAAETIINANSNPEMNPSTITQKAEQVNSS----- 2114
Qy        312 FAKQFRKFGYDGDHLLAEYQRGVGTIEASLTTISTALGAVSAGVSAAAVGSAVGTPIALL 371
Db       2115 ------KTALNGDENLAAAKQNAKTYLNTLTSITD------------------------- 2143
Qy        372 VAGVTGLISGILEASKQAMFESVANRLQGKILEWEKQNGGQNYFDKGYDSRYAAYLANNL 431
Db       2144 -AQKNNLISQITSATRVSGVDTV------------KQNA--QHLDQA-----MASLQNG- 2182
Qy        432 KFLSELNKELEAERVIAITQQRWDNNIGELAGITKLGERIKSGKAYADAFEDGKKVEAGS 491
Db       2183 -----INNE----------------------------SQVKSSEKYRDA-DTNKQQEYDN 2208
Qy        492 NIT-----LDAKTGIIDISNS--------NGKKTQALHFTSPLLTAGTESRERLTNGKYS 538
Db       2209 AITAAKAILNKSTGPNTAQNAVEAALQRVNNAK-DALNGDAKLIAAQNAAKQHL--GTLT 2265
Qy        539 YINKLKFGRVKNWQVTDGEASSKLDFSKVIQRVAETEGTDEIGLIVNAKAGNDDIFVGQG 598
Db       2266 HIT----------------TAQRNDLTNQISQATNLAGVESV------------------ 2291
Qy        599 KMNIDGGDGHDRVFYSKDGGFGNI--TVDGTSATEAGSYTVNRKVARGDIYHEVVKRQET 656
Db       2292 KQNAN----------SLDGAMGNLQTAINDKSGTLASQNFLDADEQKRNAYNQAVSAAET 2341
Qy        657 KVGKRT------ETIQYRDYELRKVGYGYQSTDNLKSVEEVIGSQFNDVFKGSKFNDIFH 710
Db       2342 ILNKQTGPNTAKTAVEQALNNVNNAKHALNGTQNLNNAKQAAITAIN------------- 2388
Qy        711 SGEGDDLLDGGAGDDRLFGGKGNDRLSGDEGDDLLDGGSGDDVLNGGAGNDVYIFRKGDG 770
Db       2389 ----------GASD---LNQKQKDALKAQ--------------ANGAQ-------RVSNA 2414
Qy        771 NDTLYDGT------GNDKLAFAD-----ANISDIMIERTKEGIIVKRNDHSGSI--NIPR 817
Db       2415 QDVQHNATELNTAMGTLKHAIADKTNTLASSKYVNADSTKQNAYTTKVTNAEHIISGTPT 2474
Qy        818 WYITSNLQNYQSNKTDHKIEQLIGKDGSYITSDQIDKILQDKKDGTVITSQELKKLADEN 877
Db       2475 VVTTPSEVTAAANQVNSAKQELNG-----------DERLREAKQNANTAIDALTQL---N 2520
Qy        878 KSQKLSASDIASSLNKL 894
```

```
                ||    :     |:|
Db       2521 TPQKAKLKEQVGQANRL 2537

RESULT 28
US-10-114-170-257
; Sequence 257, Application US/10114170
; Publication No. US20030023075A1
;   GENERAL INFORMATION:
;        APPLICANT: Blattner, Frederick R.
;                   Burland,     Valerie
;                   Perna,  Nicole T.
;                   Plunkett, Guy
;                   Welch,  Rod
;        TITLE OF INVENTION: No. US20030023075A1el Sequences of E. coli O157
;        NUMBER OF SEQUENCES: 265
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Quarles & Brady
;             STREET: 1 South Pinckney Street
;             CITY: Madison
;             STATE: WI
;             COUNTRY: US
;             ZIP: 53701-2113
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Diskette, 3.50 inch. 1.44Mb storage
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: Word Perfect 8.0
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/114,170
;             FILING DATE: 01-Apr-2002
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 09/453,702
;             FILING DATE: 03-DEC-1999
;             APPLICATION NUMBER: 60/110,955
;             FILING DATE: 04-DEC-1998
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Seay, Nicholas J.
;             REGISTRATION NUMBER: 27386
;             REFERENCE/DOCKET NUMBER: 960296.95017
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (608) 251-5000
;             TELEFAX: (608) 251-9166
;   INFORMATION FOR SEQ ID NO: 257:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 3169 amino acids
;             TYPE: amino acid
;             STRANDEDNESS: <Unknown>
;             TOPOLOGY: linear
;        MOLECULE TYPE: protein
;        SEQUENCE DESCRIPTION: SEQ ID NO: 257:
US-10-114-170-257

  Query Match             4.0%;  Score 186.5;  DB 15;  Length 3169;
  Best Local Similarity   19.4%;  Pred. No. 0.00096;
  Matches  214;  Conservative  152;  Mismatches  327;  Indels  411;  Gaps   56;

Qy         11 IQAGLNSTKSGLKNLYLAIPKDYDPQKGGTLN--------DFI------------KAAD 49
              |:  |  | | |:::|| |:      ||:||         ||              :: |
Db       1604 IKLPLVQTGSLLRDIYLTTPEKLS-TSGGSLNIMGHLVPVSFIYDIGGVINGNRISESTD 1662

Qy         50 ELGIAR-----------------LAEEPNHTETAKKSVDTVNQFLSLTQTGIAISATKLE 92
                    |                 |:||   |:   |  ||    || :     | :   |||
Db       1663 VKNKIRSLKINGDILQHYINTHYLSEE--QTQKIKDIVD----FLGIQDNTIKV---KLE 1713

Qy         93 KFL-----------------QKHSTNKLAKGLDSVENIDRKLGKA--SNVLS-------- 125
                 :                    ::|  | |:   ||    |   || |  :  :||||
Db       1714 SDIKPISEIQQPLHSILSRQKEHVKNLLSGLLDEFSNKLRKQGLSLKTNVLSVNNFKESK 1773

Qy        126 --------TLSSFLGTALAGIELDSLI-----KKG-DAAPDALAKASIDLINEIIGNLSQ 171
                        |::   |  |  :::|: :      |:| ::   :||    :.|| |  :|| :
Db       1774 INSDTVEVTVTDLQG-RLYRVDIDTRVIGLTFKEGINSLSEALEHMNIDAIMSVIGLVQY 1832

Qy        172 STQTIEAFSSQLAKLG---STISQAKGFSNIGNKLQNLNFSKTNLGLEIITGLLSGISAG 228
                        ::: |:        |  |    |   |:| |                   |:   | ||
Db       1833 ---------ARMIKMNDNISAIDHAGAVSDIKN--------------IVDKFLGGI--- 1865

Qy        229 FALADKNASTGKKVAAGFELSNQVIGNVTKAISSYVLAQRVAAGL----STTGAVAALIT 284
                                |:|   :  |       :|    |    :::||    |   |   |
Db       1866 ----------------LTLTNNRVYN-PGGVSGASLEGFISSGLEVCASRMGGTAGRYL 1907

Qy        285 SSIMLAIS-PLAFMNA------ADKFNHANALDEFAKQFRKFGYDGDHLLAEYQRGVGT- 336
              |::    |  ||   :                 ||| |    :                      ||    : |
Db       1908 SNVAKVIKLPLLDIGINIWSLYDSSLNHAKATTQI----------------EY---ISTA 1948

Qy        337 IEASLTTISTALGAVSAGVSAAAVGSAVGTPIALLVAGVTGLISGILEASKQAMFESVAN 396
              |: | ::|:|||            ::|:      |:|: :  :|
Db       1949 IDVSFSSINTAL----------SIGAIAYPPLAIAIVPIT------------------- 1978

Qy        397 RLQGKILEWEKQNGGQNYFDKGYDSRYAAYLANNLKFLSELNKE----LEAERVIAITQQ 452
                    |      | :|                  ||  |         ::::|:        ||||: :
Db       1979 -----IFSHEVKN-------------YAVY-------VNQINERHKLWLEAEKYL----- 2008

Qy        453 RWDNNIGELAGITKLGERIKSGKAYADAFEDGKKVEAGSNITLDAKTGIIDISNSN--GK 510
                 ||    ::   | |                                  |||||:||:      |
Db       2009 --DNGSAKVLSINK-------------------------------ATGIIDLSNNQVLGN 2035

Qy        511 KTQALHFTSPLLTAGTESRERLTNGKYSYINKLKFGRVKNWQVTDGEASSKLDFSKVIQR 570
                 :      |:|          :|: || :    |    :  :||  |          ::    :
Db       2036 IYLDMRENPPIL-----------HGEKSYNSGKNIG--SHPDMTDREIMESRAYNFACTK 2082

Qy        571 VAETEGTDEIGLIVNAKAGNDDIFVGQGKMNIDG---------GDGH------------- 608
              :::                || ||| | |   |           :|:
Db       2083 LSD-------------AGEPDIF-GWGDKEICNSMELSESQLANGYSNRQWPSQIPVIP 2127

Qy        609 DRVFYSKDGGFGNITVDGTSATEAGS---YTVNRKVARGDIYHEV--VKRQETKV--GKR 661
              : :: :     |:|       |     | :   :     | :          ::     :     |    |    :  |   ||
Db       2128 EGIYNTVYLGYGETLRANTEVTLSSTGYFYEIARAYTDDELSEPLLTVCNQHSHVIGGKE 2187

Qy        662 TETI------------------QYRDYELRKVG--------------YGYQSTDNLKSV- 688
               ||                  ::::|    |                 |  :  |  ::::
Db       2188 PLTIIIPAIEHGMLGSNLHMIERFKNYNFSISGGKGGIKLMVGGIGDYNIECTPGVRNII 2247

Qy        689 -EEVIGSQFN---DVFKGSKFNDIFH--SG--EGDDLLDGGAGDDRLFGGK-GNDRLSGD 739
                | :    ||    |:  | | |  || ||   |   :       | : :  | | | |:: |:
Db       2248 SFEQLSRDFNLDLDLSDGRKQNFNFHHPSGFYSGKVMSITQKGINAVVGTKLGYDKIRGN 2307

Qy        740 EGDDLLDGGSGDDVLNGGAGND-------------VYIFRKGDGNDTLYDGTGNDKLAFA 786
                |:    |||  |:  | |::                   :||  | :||   :   |   : :|:
Db       2308 NLDNTFSLGSGGGVIYSGGGSNTYFVPATLQDNLHIYISEKSNGNHIIL-GDMHSRLSI- 2365

Qy        787 DANISDIMIERTKEGIIVKRNDHSGSINIPRWYITSNLQNYQSNKTDHKIEQLIGKDGSY 846
              : : |:          |   :|   ::|     |         :::: | |
Db       2366 ECHFSN------NEREFIKMGYYNGCDVILESDTIQKIKSFAKNITIQ------------ 2407

Qy        847 ITSDQIDKILQDKKDGTVITSQELKKLADENKSQK--LSASDIASSLNKLVGSMALFG-- 902
               |:|  :       ||  :     |  ::      |:||:  |       |       :   ::  : :||
Db       2408 -TADGVMANWDDKLNTLSVYSIDMIAWRDKNKTAKEPLPVDVIQLTNWRMHNTCSLFYDK 2466

Qy        903 -----TANSVSSNALQPITQ-PTQ 920
                        || ::    | | |: | |
Db       2467 YQVDIEANKLTYTVLFPDTELPVQ 2490

RESULT 29
US-09-815-242-12996
; Sequence 12996, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
;  APPLICANT: Haselbeck, Robert
;  APPLICANT:  Ohlsen, Kari L.
;  APPLICANT:  Zyskind, Judith W.
;  APPLICANT:  Wall, Daniel
;  APPLICANT:  Trawick, John D.
;  APPLICANT:  Carr, Grant J.
```

```
;  APPLICANT:  Yamamoto, Robert T.
;  APPLICANT:  Xu, H. Howard
;  TITLE OF INVENTION: Identification of Essential Genes in
;  TITLE OF INVENTION:  Prokaryotes
;  FILE REFERENCE: ELITRA.011A
;  CURRENT APPLICATION NUMBER: US/09/815,242
;  CURRENT FILING DATE:  2001-03-21
;  PRIOR APPLICATION NUMBER: 60/191,078
;  PRIOR FILING DATE: 2000-03-21
;  PRIOR APPLICATION NUMBER: 60/206,848
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 60/207,727
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: 60/242,578
;  PRIOR FILING DATE: 2000-10-23
;  PRIOR APPLICATION NUMBER: 60/253,625
;  PRIOR FILING DATE: 2000-11-27
;  PRIOR APPLICATION NUMBER: 60/257,931
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: 60/269,308
;  PRIOR FILING DATE: 2001-02-16
;  NUMBER OF SEQ ID NOS: 14110
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 12996
;   LENGTH: 6281
;   TYPE: PRT
;   ORGANISM: Staphylococcus aureus
US-09-815-242-12996

  Query Match             4.0%;  Score 186.5;  DB 9;  Length 6281;
  Best Local Similarity   18.6%;  Pred. No. 0.0027;
  Matches  224;  Conservative  153;  Mismatches  477;  Indels  349;  Gaps   51;

Qy          1 MSNINVIKSNIQAGLNST--KSGLKNLYLAIPKDYDPQKGGTLNDFIKAADELGIARLAE 58
                : |:: |: :      | |  |:  ||      :     : ||       :|    |:  ::
Db        208 IQNVSSAKNALNGDQNVTNAKNAAKNALNNLTSINNAQKRDLTTKIDQATTVAGVEAVSN 267

Qy         59 EPNHTETAKKSV-----DTVNQFLSLT--------QTGIAISATKLEKFLQKHSTNKLAK 105
                     || ::        |  |   |              :|    :  |  |  | |:|  : |  |
Db        268 TSTQLNTAMANLQNGINDKTNTLASENYHDADSDKKTAYTQAVTNAENILNKNSGSNLDK 327

Qy        106 GLDSVENIDRKLGKASNVLS------TLSSFLGTALAGI------ELDSLIKKGDAAPDA 153
                  :|||   ::    |    |:              |    |  :  |:          :  |  |  ::    |  :
Db        328 --TAVENALSQVANAKGALNGNHNLEQAKSNANTTINGLQHLTTAQKDKLKQQVQQAQNV 385

Qy        154 ----LAKASIDLINEIIGNLSQSTQTIEAFSSQLAKLGSTISQAKGFSNIGNKLQNLN-- 207
                    |:|  :  :|  :|  |  |   |     |    :        |  :|     :   :|   |   :  :  |
Db        386 AGVDTVKSSANTLNGAMGTLRNSIQDNTATKNGQNYLDAT---ERNKTNYNNAVDSANGV 442

Qy        208 ---FSKTNLGLEIITGLLSGISAGFALADKNASTGKKVAAGFELSNQVIGNVTKAISSYV 264
                     |   |:       |   :  :  ::         :  |||   |                :  |     |         ::
Db        443 INATSNPNMDANAINQIATQVT-----STKNALDG---THNLTQAKQTATNAIDGATNLN 494

Qy        265 LAQRVAAGLSTTGAVAALITSSIMLAISPL-----AFMNAADKFNHANALDEF--AKQFR 317
                ||:  |              |       :||         :|            :    |    |      ::  |:|  :
Db        495 KAQKDALKAQVTSAQRVANVTSIQQTANELNTAMGQLQHGIDDENATKQTQKYRDAEQSK 554

Qy        318 KFGYD----GDHLLAEYQRGVGTIEAS-------LTTISTAL-GAVSAGVSAAAVGSAVG 365
                |   ||          :     |  |   :  :|:        :|:     ||   |         :  ||      :|
Db        555 KTAYDQAVAAAKAILNKQTGSNSDKAAVDRALQQVTSTKDALNGDAKLAEAKAAAKQNLG 614

Qy        366 TPIALLVAGVTGL---------ISGILEASKQA-MFESVANRLQGKILEWEKQNGGQNYF 415
                |    :    |    |  |              :  |:         |       :      |  |||  |  :  :          |||
Db        615 TLNHITNAQRTDLEGQINQATTVDGVNTVKTNANTLDGAMNSLQGSINDKDATLRNQNYL 674

Qy        416 DKGYDSRYAAYL---------------------------------------ANNLKFLSE 436
                |    :|:    ||                                                                       |:||:
Db        675 DAD-ESKRNAYTQAVTAAEGILNKQTGGNTSKADVDNALNAVTRAKAALNGADNLR---- 729

Qy        437 LNKELEAERVI----AITQQRWDN---------NIGELAGITKLGERI------------ 471
                |  :       |          |          :||   :   ||              |:    :  |:      |     :
Db        730 -NAKTSATNTIDGLPNLTQLQKDNLKHQVEQAQNVAGVNGVKDKGNTLNTAMGALRTSIQ 788

Qy        472 -----KSGKAYADAFEDGKKVEAGSNITLDAKTGIIDISNS--------NGKKTQ----- 513
                      |:  :   |   ||   :     |             |    ::         |:|:  :|:              ||      |
Db        789 NDNTTKTSQNYLDASDSNKN---NYNTAVNNANGVINATNNPNMDANAINGMANQVNTTK 845

Qy        514 -ALHFTSPLLTAGTESRERLTNG-----------KYSYINKLKFGRVKNWQVTDGEASSK 561
                ||:        |    |   |    :       :     |                          |        |      :        |  |  |      |   :|
Db        846 AALNGAQNLAQAKTNATNTINNAHDLNQKQKDALKTQVNNAQRVSDANNVQHTATELNSA 905

Qy        562 LDFSKVIQRVAETEGTDEIGLIVNA----------KAGN---------------DDIFVG 596
                  :      |          :|  :   |     |        |    |||              |     |                                       :|:
Db        906 MTALKA--AIADKERTKASGNYVNADQEKRQAYDSKVTNAENIISGTPNATLTVNDVNSA 963

Qy        597 QGKMNID----GGDGHDRVFYSKDGGFGNI---------------------TVDGTSAT 630
                    ::|                  ||  :   ||     :|:            |                                         |:||
Db        964 ASQVNAAKTALNGDNNLRV--AKEHANNTIDGLAQLNNAQKAKLKEQVQSATTLDGVQTV 1021

Qy        631 EAGSYTVN--RKVARGDIYHEVV-------------KRQE-------------------- 655
                  :   |  |:|        |      |      |     :|                                    |  |
Db       1022 KNSSQTLNTAMKGLRDSIANEATIKAGQNYTDASPNNRNEYDSAVTAAKAIINQTSNPTM 1081

Qy        656 -----TKVGKRTETIQYRDYELRKVGYGYQST-DNLKSVEEVIGSQFN------------ 697
                         |:|     :      |    :           |    :       :       :|  |   ::      :        :  |   :
Db       1082 EPNTITQVTSQVTTKEQALNGARNLAQAKTTAKNNLNNLTSINNAQKDALTRSIDGATTV 1141

Qy        698 -----DVFKGSKFNDIFHSGEGDDLLDGGAGDDRLFGGKGNDRLSGDEGDDLLDG----- 747
                      :   |  ::  |:    ||                    |    |                       ||            :          ||
Db       1142 AGVNQETAKATELNNAMHS------LQNGI----------NDETQTKQTQKYLDAEPSKK 1185

Qy        748 GSGDDVLNGGAGNDVYIFRKGDGNDTLYDGTGNDKLAFADANISDIMIERTKEGIIVKRN 807
                  :  |     :|               |       |      |     :                   ||   |       |   :  ::      :|                |    |
Db       1186 SAYDQAVNAAKA----ILTKASGQNV-------DKAAVEQA-LQNVNSTKTALNGDAKLN 1233

Qy        808 DHSGSINIPRWYITSNLQNYQSNKTDHKIEQLIGKDGSYITS----------DQIDKILQ 857
                  :        :                :|  ::  |    |           |::|    |           :|                                 |::    ::
Db       1234 EAKAAAKQTLGTLT-HINNAQRTALDNEITQATNVEGVNTVKAKAQQLDGAMGQLETSIR 1292

Qy        858 DKKDGTVITSQELKKLADENK---SQKLSASDIASSLNKLVGSMALFGTANSVSSNALQP 914
                ||       |    :    ||        :               |         :         ||    ::|:       |:      |||       |                        |     :              |:|
Db       1293 DK--DTTLQSQNYQDADDAKRTAYSQAVNAA--ATILNKTAGG----NTPKADVERAMQA 1344

Qy        915 ITQ 917
                :||
Db       1345 VTQ 1347


RESULT 30
US-10-369-493-19046
; Sequence 19046, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
;  APPLICANT: Cao, Yongwei
;  APPLICANT:  Hinkle, Gregory J.
;  APPLICANT:  Slater, Steven C.
;  APPLICANT:  Goldman, Barry S.
;  APPLICANT:  Chen, Xianfeng
;  TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
;  TITLE OF INVENTION:  PLANTS WITH IMPROVED PROPERTIES
;  FILE REFERENCE: 38-10(52052)B
;  CURRENT APPLICATION NUMBER: US/10/369,493
;  CURRENT FILING DATE:  2003-02-28
;  PRIOR APPLICATION NUMBER: US 60/360,039
;  PRIOR FILING DATE: 2002-02-21
;  NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 19046
;   LENGTH: 1136
;   TYPE: PRT
;   ORGANISM: Anabaena PCC7120
US-10-369-493-19046

  Query Match             3.9%;  Score 180.5;  DB 12;  Length 1136;
  Best Local Similarity   19.6%;  Pred. No. 0.00054;
  Matches  197;  Conservative  125;  Mismatches  354;  Indels  327;  Gaps   44;
```

```
Qy         31 KDYDPQKGGTLNDFIKAADELGIARLAEE--PNHTETAKKSVDTVNQFLSLTQTGIAISA 88
              : |  | ||        || | |       : |   |    | ||        |||||| |
Db        260 QSYQEQLGG-----IKGALERSCQGVYESLPPQLQECAKWI------FLSLTQLGEGTED 308

Qy         89 TKLEKFLQKHSTNKLAKGLDSVENIDRKLGKASNVLSTLSSFLGTALAGIELDSLIKKGD 148
              |:   :          |   ||                  |  ||:       |  |  |:: |
Db        309 TRRRIYKSDLIVKKYPAGL--------------VEQTLNVLTNAKLVVINLEAEI---- 349

Qy        149 AAPDALAKASIDLINEIIGNLSQSTQTIEAFSSQLAKLGSTI--------SQAKGFSNIG 200
                 :|  |:       :          |||     |:|       | :  ||:        :  :     |
Db        350 ---EAQGKS----FSPTPPNLSTPFVTVEVAHEILIRHWSTLRWWLEENRDRLRKQRQIN 402

Qy        201 NKLQNLNFSKTNLGLEIITGLLSGISAGFALADKNASTGKKVAAGFEL----SNQVIGNV 256
               :  |                   || | |  |        | ::|     ::      :::: :|
Db        403 HACQ-------------LWQQSGKQADFLL------QGARLAEAEDIYIYWTDELGADV 442

Qy        257 TKAISSYVL--------------AQRVAAGLSTTGAVAALITSSIMLAISPLAFMNAAD 301
               : | : :                 |||    ||  |   :          ::   ||:    :
Db        443 QEFIGACLAERKHQQLQAKNRLKQAQRAVVALSVLGIAS--------VSFGGLAYWQGRE 494

Qy        302 -KFNHANALDEFAKQFRKFGYDGDHLLAEYQRGVGTIEASLTTISTALGAVSAGVSAA-A 359
               :|     ||: ::                  :||:  |                 |  |:  |  : ||
Db        495 AQFREIAALNSSSQA---------NLLSHQQ---------------LAALIASLKAAQQ 529

Qy        360 VGSAVGTPIALLVAGVTGLISGILEASKQAMFE-SVANRLQGK---ILEWEKQNGGQNYF 415
               |    :   |   :| |  |      :||:||       |||:|      ::          ||
Db        530 VNHVIAVPNNLKLATVTTL--------QQALFEMQERNRLEGHKDGVISISISRDGQTIA 581

Qy        416 DKGYDSRYAAYLANNLKFLSELNKELEAERVIAI-------------TQQRWDNNIGEL 461
                 |      :   :   | :  ||   :|    ::                 |  :  |   :  |  |
Db        582 SGSLDKTIKLWSRDGRLFRT-LNGHEDAVYSVSFSPDGQTIASGGSDKTIKLWQTSDGTL 640

Qy        462 AGITKLGERIKSGKAYADAFEDGKKVEAGSN--------------ITLDA-KTGIIDIS 505
                 |   |          |     |||  :  :  |:               :||        |:|  :
Db        641 LK-TITGHEQTVNNVYFS--PDGKNLASASSDHSIKLWDTTSGQLLMTLTGHSAGVITVR 697

Qy        506 NSNGKKTQALHFTSPLLTAGTESRE-----------RLTNGKYSYINKLKFG------- 546
                |    :|           : ||:|  :               :  ||     ::|  |  |
Db        698 FSPDGQT---------IAAGSEDKTVKLWHRQDGKLLKTLNGHQDWVNSLSFSPDGKTLA 748

Qy        547 ------RVKNWQVTDGEASSKL---------DFSKVIQRVAETEGTDEI---------- 580
                     :| |:: ||:      |                 :||   :  :|    |
Db        749 SASADKTIKLWRIADGKLVKTLKGHNDSVWDVNFSSDGKAIASASRDNTIKLWNRHGIEL 808

Qy        581 ------------------GLIVNAK------------------AGNDDI----FVGQG 598
                                   :| :|                    |||   :    |:   |
Db        809 ETFTGHSGGVYAVNFLPDSNIIASASLDNTIRLWQRPLISPLEVLAGNSGVYAVSFLHDG 868

Qy        599 K-MNIDGGDGHDRVFYSKDGGF-----GNITVDGTSATEAGSYTVNRKVARGDIYHEVVK 652
                :       | ||: ::::|:||          ||  :  |  |  |   |     :      |      |     : ||
Db        869 SIIATAGADGNIQLWHSQDGSLLKTLPGNKAIYGISFTPQGDLIAS---ANAD---KTVK 922

Qy        653 RQETKVGKRTETIQYRDYELRKVGY-------GYQSTDNLKSVEEVIGSQFNDVFKGSKF 705
                 :  ||  :|:   | |:  || :              |||    :    |    :|       ||
Db        923 IWRVRDGKALKTLIGHDNEVNKVNFSPDGKTLASASRDNTVKLWNVSDGKFKKTLKGHT- 981

Qy        706 NDIFHSGEGDDLLDGGAGDDRLFGGKGNDRLSGDEGDDLLDGGSGDDVLNGGAGNDVYIF 765
              :::|          |                      ||          |  |:     |  |  ||: :  :  |  ||:
Db        982 DEVFWVSFSPD------------GKIIASASADKTIRLWDSFSGNLIKSLPAHNDLVYS 1028

Qy        766 RKGDGNDTLYDGTGNDKLAFADANISDIMIERTKEGIIVKRNDHSGSINIPRWYITSNLQ 825
                : : ::     |   ||                   :   :  |: :|  ::   :    ||     |:
Db       1029 VNFNPDGSMLASTSADK---------TVKLWRSHDGHLL--HTFSGHSNVV--------- 1068

Qy        826 NYQSNKTDHKIEQLIGKDGSYITSDQIDKILQD-KKDGTVITS 867
                | |:             || || |    || ::   : || ::|:
Db       1069 -YSSS---------FSPDGRYIASASEDKTVKIWQIDGHLLTT 1101
```

RESULT 31
US-10-369-493-20095
; Sequence 20095, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 20095
; LENGTH: 210
; TYPE: PRT
; ORGANISM: No. US20030233675A1toc punctiforme
US-10-369-493-20095

```
  Query Match             3.9%;  Score 180;  DB 12;  Length 210;
  Best Local Similarity   24.1%;  Pred. No. 4.5e-05;
  Matches   68;  Conservative   45;  Mismatches   81;  Indels   88;  Gaps   14;

Qy        575 EGTDEIGLIVNAKAGNDDIFVGQGKMNIDGGDGHDRVFYSKDGGFGNITVDGTSATEAGS 634
              |||     |    :        |  |  :         ::   |||:|::|   ::|     |  :  ||  :
Db          1 EGTK--GNTIQTGNGEDTVLAGSDS-SVSTGDGNDQIFIGQNGPAQNTSADGGNG----- 52

Qy        635 YTVNRKVARGDIYHEVVKRQETKVGKRTETIQYRDYELRKVGYGYQSTDNLKSVEEVIGS 694
                            ::|:   |                                   ::||      :|
Db         53 ------------NDVITVVEA---------------------SGSNNL------LGG 70

Qy        695 QFND---VFKGSKFNDIFHSGEGDDLLDGGAGDDRLFGGKGNDRLSGDEGDDLLDGGSGD 751
                 :|   |  :||:    :|   |  |:|  |    |   ::||:||  |:|:|    :   |   |  ||  ||
Db         71 TGDDSLTVIEGSR-QSLF-GGSGNDTLRSGGSNNRLYGGSGDDKLFSNFNDS-LSGGDGD 127

Qy        752 DVLNGGAGNDVYIFRKGDGNDTLYDGTGNDKLAFADAN-------ISDIMIERTKEGIIV 804
              |||  |                | :  |   ||| |:     |:||          ::|  ::    |  |:
Db        128 DVLFAGQ----------QGGNRLTGGTGADQFWIANANLPTSKNIVTDFVVGIDKIGL-- 175

Qy        805 KRNDHSGSINIPRWYITSNLQNYQSNKT-----DHKIEQLIG 841
                    | |  : ::   :  ||   |  :  |        :  ::   |:|
Db        176 ------GGIGVTQFSALTLLQ--QGSDTLVKAGNTELASLVG 209
```

RESULT 32
US-10-369-493-20096
; Sequence 20096, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 20096
; LENGTH: 273
; TYPE: PRT
; ORGANISM: No. US20030233675A1toc punctiforme
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(273)
; OTHER INFORMATION: unsure at all Xaa locations
US-10-369-493-20096

```
  Query Match             3.9%;  Score 180;  DB 12;  Length 273;
  Best Local Similarity   25.8%;  Pred. No. 6.7e-05;
  Matches   88;  Conservative   37;  Mismatches  100;  Indels  116;  Gaps   19;

Qy        459 GELAG--ITK-LGERIKSGKAYADAFEDGKKVEAGSNITLDAKTGIIDISNSNGKKTQAL 515
Db         30 GDLKGRAITEALGGRPGCSSLYXRSI-DFEKTTNGQATVLDSQS----VGNTQ------- 77

Qy        516 HFTSPLLTAGTESRERLTNGKYSYINKLKFGRVKNWQVTDGEASSKLDFSKVIQRVAETE 575
Db         78 YFTA----AGFDGR---TVGLTTL--DLQFGSTNSDDVT----------LKPNQTLFAGD 118

Qy        576 GTDEI----GLIVNAKAGNDDIFVGQGKMNIDGGDGHDRVFY---------SKDGGFGNI 622
Db        119 GADFVEGTKGNTIVTGNGEDTVLVGSGS-SVSTGDGNDQVFIGVNSPASNTSADGGNGN- 176

Qy        623 TVDGTSATEA-GSYTVNRKVARGDIYHEVVKRQETKVGKRTETIQYRDYELRKVGYGYQS 681
Db        177 --DEVTVVEANGSNNL-------------------------------------FGGAG 195

Qy        682 TDNLKSVEEVIGSQFNDVFKGSKFNDIFHSGEGDDLLDGGAGDDRLFGGKGNDRLSGDEG 741
Db        196 ADTLTVVE-----------GSR--QLSFGGSGNDTLKSNGSNNRLYGGSGDDKLFSSVN 241

Qy        742 DDLLDGGSGDDVLNGGAGNDVYIFRKGDGNDTLYDGTGNDK 782
Db        242 DSLF-GGDGDDVLFAGQ----------QGGNRLTGGTGADQ 271

RESULT 33
US-09-841-786-6
; Sequence 6, Application US/09841786
; Patent No. US20020054883A1
; GENERAL INFORMATION:
;  APPLICANT: NAGARAJA, T. G.
;  APPLICANT:  STEWART, GEORGE C.
;  APPLICANT:  NARAYANAN, SANJEEV K.
;  APPLICANT:  CHENGAPPA, M. M.
;  TITLE OF INVENTION: RECOMBINANT FUSOBACTERIUM NECROPHORUM LEUKOTOXIN
;  TITLE OF INVENTION:  VACCINE AND PREPARATION THEREOF
;  FILE REFERENCE: 30296
;  CURRENT APPLICATION NUMBER: US/09/841,786
;  CURRENT FILING DATE:  2001-04-24
;  PRIOR APPLICATION NUMBER: 09/558,257
;  PRIOR FILING DATE: 2000-04-25
;  NUMBER OF SEQ ID NOS: 15
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
;   LENGTH: 773
;   TYPE: PRT
;   ORGANISM: Fusobacterium necrophorum
US-09-841-786-6

  Query Match             3.9%;  Score 179;  DB 9;  Length 773;
  Best Local Similarity   21.4%;  Pred. No. 0.00039;
  Matches  172;  Conservative  114;  Mismatches  310;  Indels  208;  Gaps   38;

Qy        111 ENIDRKLGKASNVLST-------LSSFLGT-ALAGIELDSLIKKGDAAPDALAKAS--ID 160
Db         14 KNVDITTEDKNNITSTGGLGTAGLASASGTVAVTNIKRNS----GVTVENSFVKAAEKVN 69

Qy        161 LINEIIGNLSQSTQTIEAFSSQLAKLGSTISQAK----GFSNI---GNKLQNLNF----- 208
Db         70 VRSDITGNVA-----LTAYQGPVGALGIGAAYAELNSNGRSNISIKNSKLLGKNIDVIVK 124

Qy        209 SKTNLGLEIITGLLSGISAGFALADK-------------------NASTGKKVAAGFELS 249
Db        125 DKSELRAE-AKGLTVGAVAAGAIISKAKNEMNSEVEIEKSIFNEENRVTSPSKGIGREI- 182

Qy        250 NQVIGNVTKAISSYVLAQRVAAGLSTTGAVA-ALITSSIMLAISPLAFMNAADKFN---- 304
Db        183 -----NVKVEKENRVTAESQGA---SVGAVAGAGIISEAKDAGS--SYLKVSTKSGRSIF 232

Qy        305 HANALDEFAKQFRKFGYDGDHLLAEYQRGVGTIEASLTTISTALGAVSAG-----VSAAA 359
Db        233 HADNVNMEATHKMKVTAVSKAVTGSVLGGVGVTKAEATAAGKTMVEVEEGNLFRTNRLNA 292

Qy        360 VGSAVG------TPIALLVAGVTGLI--SGILEASKQAMFESVANRLQGKILEWEKQNGG 411
Db        293 ISKVEGLDEDKVTAKSSVVSGNGGGIAGAGVNTSTAQSNTESVV-RLR-------KQDYE 344

Qy        412 QNYFDKGYDSRYAAYLANNLKFLSELNKELEAERVIAITQQRWDNNIGEL--AGITKLGE 469
Db        345 NNDYTK--------------KYISEVN-------ALALNDTKNEANIESLAVAGVHAQG- 382

Qy        470 RIKSGKAYADAFEDGKKVEAGSNIT-LDAKTGIIDISNSNGKKTQALHFTSPLLTAGTES 528
Db        383 ---TNKAFTRSNKLTSTTVNGGNVSQLRAKALAKNENYGNVKGT-----GGALVGAETAA 434

Qy        529 RERLTNGKYSYINKLKFGRVKNWQVTDGEASSKLDFSKVIQRVAETEGTDEIGLIV---- 584
Db        435 VENYTK---STTGALVAG---NWEIGD-----KLETIARDNTIVRVNGDGTKGGLVGKNG 483

Qy        585 ---------NAKAGNDD--IFVGQGKMNIDGGDGHDRVFYSKDGGFGNITVDGTSATEAG 633
Db        484 ISVKNTISGETKSSIEDKARIVGTGSVNVDALNELDVDLQGKSGGYGGIGIGNV------ 537

Qy        634 SYTVNRKVARGDIYHEVVKRQETKVGKRTETIQYRDYELRKVGYGYQSTDNLKSVEEVIG 693
Db        538 -----------DVNNVIKKNVEAKIGRH-----------------------AIVETTG 561

Qy        694 SQFNDVFKGSKFNDIFHSGEGDDLLDGGAGDDRLFGGKGNDRLSGD-EGDDLLDGGSGDD 752
Db        562 KQEYQAFTRAKVNIL-----------GKGDAAAAAAISNVHISNEMDIKNLAKQYASSQ 609

Qy        753 VLNGGAGNDVYIFRKGDGNDTLYDGTGNDKLAFADANIS-DIMIERTKEGIIVKRNDHSG 811
Db        610 LITKNSKNNITLASSSESNVNVH-GVAEARGAGAKATVSVKNQINRTNNVDLAGKIKTEG 668

Qy        812 SINIPRWYITSNLQNYQSNKTDHK 835
Db        669 NINVYAGYD----KNYNISKTNSK 688

RESULT 34
US-10-193-764-57
; Sequence 57, Application US/10193764
; Publication No. US20030133943A1
; GENERAL INFORMATION:
;  APPLICANT: Loosmore, Sheena M.
;  APPLICANT:  Yang, Yan-Ping
;  APPLICANT:  Klein, Michel H.
;  TITLE OF INVENTION: PROTECTIVE RECOMBINANT HAEMOPHILUS INFLUENZAE HIGH
;  TITLE OF INVENTION:  MOLECULAR WEIGHT PROTEINS
;  FILE REFERENCE: 1038-1239MIS
;  CURRENT APPLICATION NUMBER: US/10/193,764
;  CURRENT FILING DATE:  2002-07-12
;  PRIOR APPLICATION NUMBER: 09/167,568
;  PRIOR FILING DATE: 1998-10-07
;  NUMBER OF SEQ ID NOS: 91
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 57
;   LENGTH: 992
;   TYPE: PRT
;   ORGANISM: Haemophilus influenzae
US-10-193-764-57

  Query Match             3.9%;  Score 179;  DB 12;  Length 992;
  Best Local Similarity   22.2%;  Pred. No. 0.00056;
  Matches  207;  Conservative  110;  Mismatches  329;  Indels  286;  Gaps   54;

Qy        136 AGIELDSLIKKGDA---APDALAKASIDLINEIIGN-LSQSTQTIEAFSSQLAKLGSTIS 191
Db         16 AGI--DSEFPGGSGTKESPKTNGEQPTVLTNETISNYLKSGTWVMNITAKKNLTVNSSI- 72

Qy        192 QAKGFSNIGNKLQNLNFS--KTNLGLEI---IT---GLLSGISAGFALADKNASTG---- 239
```

```
                 |||:    :  |  | | |::|  ||   | |:  | |:    || : |
Db         73 ------NIGDSSHLILHSEGKNNGGVKIKEDITSNGGNLTIQSGGWVDVHKNITLGTGTL 126

Qy        240 ---KKVAAGFELSNQVIGNVTKAISSYVLAQRVAAG-LSTTGAVAALITSSI-------- 287
                  | :  ||      || |:   :   ||  | | :: ||     |   :::
Db        127 NITAKGSIAFE------GNGTEKARNASSAQITAQGTITNTGDQKQLRLNNVSINGTGIG 180

Qy        288 --MLAISP------------------------LAFMNAADK----FNHANALDEFAKQF 316
                ::| |                           |:|   :|:    :|    :: |  |
Db        181 LNFVSIQPNTSHRFDGELIISGRVHVNQTTPKNLSFWKVSDESYWNVSHLTVKEKSAFSF 240

Qy        317 RKFGYDGDH--LLAEYQRGVGTIEASLTTISTALGAVSAGVSAAAVGSAVGTPIALLVAG 374
                ||  : :|     : |::| | |  |      |   :    :| :
Db        241 TKFALNNNHGRETSRYRKGGGVIFRSPT------GHTNFTVKQGSVANF----------- 283

Qy        375 VTGLISGILEASKQAMFESVANRLQ-----------GKILEWEKQNGGQNYFDKGYDSR 422
                         : |      : ||:|          ||:|      |          |  |
Db        284 ----------SFKAKNDTNHANQLPIQFNSNISVDGGGKVLFCITSNYSGRSVGIGMSSI 333

Qy        423 YAAYLANNLKFLSELNKELEAERVIAITQQRWDNNIGELAGITKLGERIKSGKAYADAFE 482
                  :   :|| |     |   :   :    |::   |  |         :||:      |:| |
Db        334 NVSD-GSNLTF----NSSIRGQEAFNISK---DLTINATGSFFELGQ-------YSDTFN 378

Qy        483 -DGKKVEA----------GSNITLDAK------TGIIDIS----------NSNGKKTQAL 515
                :|    :|            | |:||   :        || |:||           | ||:   |
Db        379 GNGFNHDAIKSTHNISILGGNVTLGGQDSSSTITGNINISQAANVTLRAYNGNGRNKQ-- 436

Qy        516 HFTSPLLTAGTESRE---------RLTNGKYSYINKLKF-GRVK-NWQVTD---GEASSK 561
                       || |   ||                ||  |     | |  |   :   :|          ||
Db        437 ------LTLGNVSIEGNLSLIGASANINGNLSVKENAKFKGETQDNLNITGTFINNGDSK 490

Qy        562 LDFSKVIQRVAETEGTDEIGLIVNAKAGNDDIFVG-----QGKMNIDGGDGHDRV----- 611
              :: |: : ::       :: : :||       :| :||     : : :
Db        491 INISQGVVKLGNVTNDGDLNITTHAKHNQRSIIGGDIINKKGSLNITDSNKNAEIQIGGN 550

Qy        612 FYSKDGGFG------NIT--------VDG----TSATEAGSYTVNRKVARGDIYHEVVKR 653
                  |:|         |||          |||    : ||    :|:  |           |:
Db        551 ISQKEGNLTISSDKINITNQITIKAGVDGENSDSDATNNANLTIKTK--------ELKLT 602

Qy        654 QETKV-GKRTETIQYRDYELRKVGYGYQSTD--NLKSVEEVIGSQFNDVFKGSKFNDIFH 710
              |: : |      | :|     :|   | |  | | |          || | | || :
Db        603 QDLNISGFNKAEITAKDGSDLTIG-NTNSADSTNAKKV------TFNQV-KDSKIS---- 650

Qy        711 SGEGDDLLDGGAGDDRLFGGKGNDRLSGDEGDDLLDGGSGDDVLNGG---AGNDVYIFRK 767
                                  |||  :     | ::      |||:  |  |    |  :| :
Db        651 -----------AGDHNV---TLNSKVETSGNTDNTGDGSGN---NAGLTIAAKNVEVKNN 693

Qy        768 GDGNDTLYDGTGNDKLAF-ADANISDIM--IERT-KEGII---VKRNDHSGSINIPRWYI 820
                | |: : | ::||  |||  |:     :| | | |  |    ||   ||::||
Db        694 ITSNKTV-NITASEKLTTKADATINATTGNVEVTAKTGDIKGEVKST--SGNVNITANGD 750

Qy        821 TSNLQNYQSNKTDHKIEQLIGKDGSYITSDQIDKILQDKKDGTVITSQELKKLADENKSQ 880
               | |: |   |             ||:|:               :|:|    :    |
Db        751 TLNVSNVSGNAVT-------------ITADK-----------GKLTTQASSSITSNNGQT 786

Qy        881 KLSASD--IASSLNKLVGSMALFGTANSVSSN 910
               |:| |  || |:|    ::   ||  :| :
Db        787 TLTAKDGSIAGSINAANVTLNTTGTLTTVEGS 818
```

RESULT 35
US-10-193-764-55
; Sequence 55, Application US/10193764
; Publication No. US20030133943A1
; GENERAL INFORMATION:
; APPLICANT: Loosmore, Sheena M.
; APPLICANT: Yang, Yan-Ping
; APPLICANT: Klein, Michel H.
; TITLE OF INVENTION: PROTECTIVE RECOMBINANT HAEMOPHILUS INFLUENZAE HIGH
; TITLE OF INVENTION: MOLECULAR WEIGHT PROTEINS
; FILE REFERENCE: 1038-1239MIS
; CURRENT APPLICATION NUMBER: US/10/193,764
; CURRENT FILING DATE: 2002-07-12
; PRIOR APPLICATION NUMBER: 09/167,568
; PRIOR FILING DATE: 1998-10-07
; NUMBER OF SEQ ID NOS: 91
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 55
; LENGTH: 998
; TYPE: PRT
; ORGANISM: Haemophilus influenzae
US-10-193-764-55

```
  Query Match             3.9%;  Score 179;  DB 12;  Length 998;
  Best Local Similarity  22.2%;  Pred. No. 0.00057;
  Matches  207;  Conservative  110;  Mismatches  329;  Indels  286;  Gaps   54;

Qy        136 AGIELDSLIKKGDA---APDALAKASIDLINEIIGN-LSQSTQTIEAFSSQLAKLGSTIS 191
              |||  ||    |     :|    :      | || | | |    |  :  :   : :  |:|
Db         22 AGI--DSEFPGGSGTKESPKTNGEQPTVLTNETISNYLKSGTWVMNITAKKNLTVNSSI- 78

Qy        192 QAKGFSNIGNKLQNLNFS--KTNLGLEI---IT---GLLSGISAGFALADKNASTG---- 239
                    |||:   :  |  | | |::|  ||   | |:  | |:    || : |
Db         79 ------NIGDSSHLILHSEGKNNGGVKIKEDITSNGGNLTIQSGGWVDVHKNITLGTGTL 132

Qy        240 ---KKVAAGFELSNQVIGNVTKAISSYVLAQRVAAG-LSTTGAVAALITSSI-------- 287
                  | :  ||      || |:   :   ||  | | :: ||     |   :::
Db        133 NITAKGSIAFE------GNGTEKARNASSAQITAQGTITNTGDQKQLRLNNVSINGTGIG 186

Qy        288 --MLAISP------------------------LAFMNAADK----FNHANALDEFAKQF 316
                ::| |                           |:|   :|:    :|    :: |  |
Db        187 LNFVSIQPNTSHRFDGELIISGRVHVNQTTPKNLSFWKVSDESYWNVSHLTVKEKSAFSF 246

Qy        317 RKFGYDGDH--LLAEYQRGVGTIEASLTTISTALGAVSAGVSAAAVGSAVGTPIALLVAG 374
                ||  : :|     : |::| | |  |      |   :    :| :
Db        247 TKFALNNNHGRETSRYRKGGGVIFRSPT------GHTNFTVKQGSVANF----------- 289

Qy        375 VTGLISGILEASKQAMFESVANRLQ-----------GKILEWEKQNGGQNYFDKGYDSR 422
                         : |      : ||:|          ||:|      |          |  |
Db        290 ----------SFKAKNDTNHANQLPIQFNSNISVDGGGKVLFCITSNYSGRSVGIGMSSI 339

Qy        423 YAAYLANNLKFLSELNKELEAERVIAITQQRWDNNIGELAGITKLGERIKSGKAYADAFE 482
                  :   :|| |     |   :   :    |::   |  |         :||:      |:| |
Db        340 NVSD-GSNLTF----NSSIRGQEAFNISK---DLTINATGSFFELGQ-------YSDTFN 384

Qy        483 -DGKKVEA----------GSNITLDAK------TGIIDIS----------NSNGKKTQAL 515
                :|    :|            | |:||   :        || |:||           | ||:   |
Db        385 GNGFNHDAIKSTHNISILGGNVTLGGQDSSSTITGNINISQAANVTLRAYNGNGRNKQ-- 442

Qy        516 HFTSPLLTAGTESRE---------RLTNGKYSYINKLKF-GRVK-NWQVTD---GEASSK 561
                       || |   ||                ||  |     | |  |   :   :|          ||
Db        443 ------LTLGNVSIEGNLSLIGASANINGNLSVKENAKFKGETQDNLNITGTFINNGDSK 496

Qy        562 LDFSKVIQRVAETEGTDEIGLIVNAKAGNDDIFVG-----QGKMNIDGGDGHDRV----- 611
              :: |: : ::       :: : :||       :| :||     : : :
Db        497 INISQGVVKLGNVTNDGDLNITTHAKHNQRSIIGGDIINKKGSLNITDSNKNAEIQIGGN 556

Qy        612 FYSKDGGFG------NIT--------VDG----TSATEAGSYTVNRKVARGDIYHEVVKR 653
                  |:|         |||          |||    : ||    :|:  |           |:
Db        557 ISQKEGNLTISSDKINITNQITIKAGVDGENSDSDATNNANLTIKTK--------ELKLT 608

Qy        654 QETKV-GKRTETIQYRDYELRKVGYGYQSTD--NLKSVEEVIGSQFNDVFKGSKFNDIFH 710
              |: : |      | :|     :|   | |  | | |          || | | || :
Db        609 QDLNISGFNKAEITAKDGSDLTIG-NTNSADSTNAKKV------TFNQV-KDSKIS---- 656

Qy        711 SGEGDDLLDGGAGDDRLFGGKGNDRLSGDEGDDLLDGGSGDDVLNGG---AGNDVYIFRK 767
                                  |||  :     | ::      |||:  |  |    |  :| :
Db        657 -----------AGDHNV---TLNSKVETSGNTDNTGDGSGN---NAGLTIAAKNVEVKNN 699

Qy        768 GDGNDTLYDGTGNDKLAF-ADANISDIM--IERT-KEGII---VKRNDHSGSINIPRWYI 820
                | |: : | ::||  |||  |:     :| | | |  |    ||   ||::||
Db        700 ITSNKTV-NITASEKLTTKADATINATTGNVEVTAKTGDIKGEVKST--SGNVNITANGD 756

Qy        821 TSNLQNYQSNKTDHKIEQLIGKDGSYITSDQIDKILQDKKDGTVITSQELKKLADENKSQ 880
               | |: |   |             ||:|:               :|:|    :    |
Db        757 TLNVSNVSGNAVT-------------ITADK-----------GKLTTQASSSITSNNGQT 792
```

```
Qy        881 KLSASD--IASSLNKLVGSMALFGTANSVSSN 910
Db        793 TLTAKDGSIAGSINAANVTLNTTGTLTTVEGS 824

RESULT 36
US-09-815-242-5835
; Sequence 5835, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
;  APPLICANT: Haselbeck, Robert
;  APPLICANT:  Ohlsen, Kari L.
;  APPLICANT:  Zyskind, Judith W.
;  APPLICANT:  Wall, Daniel
;  APPLICANT:  Trawick, John D.
;  APPLICANT:  Carr, Grant J.
;  APPLICANT:  Yamamoto, Robert T.
;  APPLICANT:  Xu, H. Howard
;  TITLE OF INVENTION: Identification of Essential Genes in
;  TITLE OF INVENTION:  Prokaryotes
;  FILE REFERENCE: ELITRA.011A
;  CURRENT APPLICATION NUMBER: US/09/815,242
;  CURRENT FILING DATE:  2001-03-21
;  PRIOR APPLICATION NUMBER: 60/191,078
;  PRIOR FILING DATE: 2000-03-21
;  PRIOR APPLICATION NUMBER: 60/206,848
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 60/207,727
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: 60/242,578
;  PRIOR FILING DATE: 2000-10-23
;  PRIOR APPLICATION NUMBER: 60/253,625
;  PRIOR FILING DATE: 2000-11-27
;  PRIOR APPLICATION NUMBER: 60/257,931
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: 60/269,308
;  PRIOR FILING DATE: 2001-02-16
;  NUMBER OF SEQ ID NOS: 14110
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5835
;   LENGTH: 2434
;   TYPE: PRT
;   ORGANISM: Staphylococcus aureus
US-09-815-242-5835

  Query Match             3.8%;  Score 177.5;  DB 9;  Length 2434;
  Best Local Similarity   17.9%;  Pred. No. 0.0028;
  Matches  229;  Conservative  164;  Mismatches  450;  Indels  433;  Gaps   51;

Qy          2 SNINVIKSNIQAGLN---STKSGLKNLYLAIPKDYDPQKGGTLNDFIKAADEL------- 51
Db       1003 SELNTAMSNLQNGINDEAATKAAQKY------TDADREKQTAYNDAVTAAKTLLDKTAGS 1056

Qy         52 ---------------------GIARLAEEPNHTETAKKSVDTVN-----QFLSLTQ--- 81
Db       1057 NDNKAAVEQALQRVNTAKTALNGDERLNEAKN---TAKQQVATMSHLTDAQKANLTSQIE 1113

Qy         82 --------TGIAISATKLEKFL--------QKHSTNKL-------------------AK 105
Db       1114 SGTTVAGVQGIQANAGTLDQAMNQLRQSIASKDATKSSEDYQDANADLQNAYNDAVTNAE 1173

Qy        106 GLDSVEN--------IDRKLGKASNVLSTLS-------------SFLG--TALAGIELDS 142
Db       1174 GIISATNNPEMNPDTINQKASQVNSAKSALNGDEKLAAAKQTAKSDIGRLTDLNNAQRTA 1233

Qy        143 LIKKGDAAPD----ALAKASIDLINEIIGNL----------SQSTQTIEAFSSQLAKLGS 188
Db       1234 ANAEVDQAPNLAAVTAAKNKATSLNTAMGNLKHALAEKDNTKRSVNYTDADQPKQQAYDT 1293

Qy        189 TISQAKGFSNIGNKLQNLNFSKTNLGLEIITGLLSGISAGFALADKNASTGKKVAAGFEL 248
Db       1294 AVTQAEAITNANG--SNANETQVQAALNQLNQAKNDLNGDNKVAQAKESAKRALASYSNL 1351

Qy        249 SN-------QVIGNVTKAISSYVLAQRVAAGLSTTGAVAALITSSIMLAISPLAFMNAAD 301
Db       1352 NNAQSTAATSQIDNAT-TVAGVTAAQNTANELNT------------------AMGQLQN 1391

Qy        302 KFNHANALDE-----FAKQFRKFGYDGDHLLAEYQRGV-----------GTIEASLTTIS 345
Db       1392 GINDQNTVKQQVNFTDADQGKKDAYTNAVTNAQ---GILDKAHGQNMTKAQVEAALNQVT 1448

Qy        346 TALGAVSAGVSAAAVGSAVGTPIALL--------------VAGVTGL--ISGILEASKQA 389
Db       1449 TAKNALNGDANVRQAKSDAKANLGTLTHLNNAQKQDLTSQIEGATTVNGVNGV--KTKAQ 1506

Qy        390 MFESVANRLQGKILEWEKQNGGQNYFDKGYDSRYA---------AYL-----ANNLKFLS 435
Db       1507 DLDGAMQRLQSAIANKDQTKASENYIDADPTKKTAFDNAITQAESYLNKDHGANKDKQAV 1566

Qy        436 E------------LNKELEAERVIAITQQRWDN----------------NIGE-LAGITK 466
Db       1567 EQAIQSVTSTENALNGDANLQRAKTEAIQAIDNLTHLNTPQKTALKQQVNAAQRVSGVTD 1626

Qy        467 LGER--------------------IKSGK----------AYADAFEDGKKVEAGS-NIT 494
Db       1627 LKNSATSLNNAMDQLKQAIADHDTIVASGNYTNASPDKQGAYTDAYNAAKNIVNGSPNVI 1686

Qy        495 LDAKTGIIDISNSNGKKTQALHFTSPLLTAGTESRERLTNGKYSYINKLKFGRVKNWQVT 554
Db       1687 TNAADVTAATQRVNNAET-GLNGDTNLATAKQQAKDALR------------------QMT 1727

Qy        555 DGEASSKLDFSKVIQRVAETEGTDEI-----------GLIVNAKAGNDDIFVGQGKMNID 603
Db       1728 HLSDAQKQSITGQIDSATQVTGVQSVKDNATNLDNAMNQLRNSIANKDDVKASQ--PYVD 1785

Qy        604 GGDGHDRVFYSKDGGFGNI-------TVDGTSATEAGSYTVNRKVARGDIYHEVVKRQET 656
Db       1786 ADRDKQNAYNTAVTNAENIINATSQPTLDPSAVTQAANQVSTNKTALNGAQNLANKKQET 1845

Qy        657 --------------------------------KVGKRTETIQYRDYELRKVGYGYQSTDN 684
Db       1846 TANINQLSHLNNAQKQDLNTQVTNAPNISTVNQVKTKAEQL---DQAMERLINGIQDKDQ 1902

Qy        685 LKSVEEVIGSQFNDVFKGSKFN-------DIFHSGEGDDLLDGGA-------------- 722
Db       1903 VK---QSVNFTDADPEKQTAYNNAVTAAENIINQANGTNANQSQVEAALSTVTTTKQALN 1959

Qy        723 GDDRLFGGKGNDRLSGDEGDDLLDGGSG--------------------DDVLNGGAGND 761
Db       1960 GDRKVTDAKNNANQTLSTLDNLNNAQKGAVTGNINQAHTVAEVTQAIQTAQELNTAMGN- 2018

Qy        762 VYIFRKGDGNDTLYDGT-GNDKLAFADANISDIMIE--RTKEGIIVKRNDHSGSINIPRW 818
Db       2019 ----LKNSLNDK--DTTLGSQNFADADPEKKNAYNEAVHNAENILNK----STGTNVPKD 2068

Qy        819 YITSNLQNYQSNKTDHKIEQLIGK---------DG-SYITSDQIDKILQDKKDGTVITSQ 868
Db       2069 QVEAAMNQVNATKAALNGTQNLEKAKQHANTAIDGLSHLTNAQKEALKQLVQQSTTVA-- 2126

Qy        869 ELKKLADENKSQKLSASDIASSLNKLVGSMALFGTA-----------------NSVSSNA 911
Db       2127 -------EAQGNEQKANNVDAAMDKLRQSIADNATTKQNQNYTDASQNKKDAYNNAVTTA 2179

Qy        912 LQPITQPTQGILAPSV 927
Db       2180 QGIIDQTTSPTLDPTV 2195

RESULT 37
US-09-971-536-68
; Sequence 68, Application US/09971536
; Patent No. US20020159976A1
; GENERAL INFORMATION:
;  APPLICANT: Glenn, Matthew
;  APPLICANT:  Havukkala, Ilkka
```

```
;  APPLICANT:  Bloksberg, Leonard
;  APPLICANT:  Lubbers, Mark
;  APPLICANT:  Dekker, James
;  APPLICANT:  Christensson, Anna
;  APPLICANT:  Holland, Ross
;  APPLICANT:  O'Toole, Paul
;  APPLICANT:  Reid, Julian
;  APPLICANT:  Coolbear, Timothy
;  TITLE OF INVENTION: Lactobacillus rhamnosus Polynucleotides, Polypeptides and Methods
;  TITLE OF INVENTION:  Using Them
;  FILE REFERENCE: 1043c2
;  CURRENT APPLICATION NUMBER: US/09/971,536
;  CURRENT FILING DATE:  2001-10-02
;  PRIOR APPLICATION NUMBER: U.S. No. US20020159976A1 09/634,238
;  PRIOR FILING DATE: 2000-08-08
;  PRIOR APPLICATION NUMBER: U.S. No. US20020159976A1 09/724,623
;  PRIOR FILING DATE: 2000-11-28
;  PRIOR APPLICATION NUMBER: PCT/NZ01/00160
;  PRIOR FILING DATE: 2001-08-08
;  NUMBER OF SEQ ID NOS: 83
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 68
;   LENGTH: 1741
;   TYPE: PRT
;   ORGANISM: Lactobacillus rhamnosus
US-09-971-536-68

  Query Match              3.8%;  Score 176.5;  DB 10;  Length 1741;
  Best Local Similarity   19.4%;  Pred. No. 0.002;
  Matches  212;  Conservative  156;  Mismatches  394;  Indels  331;  Gaps   54;

Qy          5 NVIKSNIQAGLNSTKSGLKNLYLAIPKDYDPQKGGTLNDFIKAADELGIARLAEEPNHTE 64
              | |||     : |:  |  |  ::      :||   : : : |: |:         | :
Db         56 NAIKS------DGTEVPLSELTWSV----TDEKG----NVVASKDKNGVV-----TGHVD 96

Qy         65 TAKKSVDTVNQFLSLTQTGIAISATKLEKFLQKHSTNKLAKGLDSVENIDRKLGKASNVL 124
              :: :    |       | |  | | :    : |: : :    |  :|          | :
Db         97 NSQPTTYVV------TYTYGAESGSAKINYKQRLAASYALTGTQTV------TATGSPIT 144

Qy        125 STLSSF---LGTAL-AG-IELDSL-IKKGDAAP---DALAKASIDLINEIIGNLSQSTQT 175
               :| |   ||    || :||  |       |  |   ||| |   :             | :
Db        145 VDVSQFALSLGDGFDAGKLELSDLNFFDADGKPVAADALIKTGV-----------YSVEL 193

Qy        176 IEAFSSQLAKLGSTISQ-AKGFSNIGNKLQNLNFSKTNLGLEIITGLLSGISAGFALADK 234
              ||  :::||| :   | | |:      |        |      || ||    :||     |
Db        194 SEAAWARIAKLTNDEGQSAAGYDFTGTSTAQLIIGLT------ATGHLS--DSGFVYDGK 245

Qy        235 -NASTGKKVAAGFELSNQVIGNVTKAISSYVLAQRVAAGLST---------TGAVAALIT 284
               || | :|    ||:       :       : | :: :| |            : ||:
Db        246 TTASQSKDLAVTVTLSDGTQKEMNLTSEDFSLVEKDSANVGTYHYLLNSVGFARLQALLG 305

Qy        285 SSIMLAISPLAFMNAADKFNHANALDEFAKQFRKFGYDGDHLLAEYQRGV---------- 334
               ::   |    |    : |     |              | |||    :|  :|:
Db        306 DTV--TIDQTAINQNSGKITITPAPATVNSNSTDFEYDGKTKASE-AKGIQATVKLGETG 362

Qy        335 GTIEASLTTISTALGAVSAG--------VSAAAVGSAVGTPIALL---VAGVTGLIS--- 380
               ||: :    |      | ||            | : :|  |     |   :| ||| |:
Db        363 KTIDLTSADIVVENDGVDAGKYSYELSDAGKAKLQAATGNNYQLTADDLAKVTGAITITP 422

Qy        381 ---------------GILEASKQAMFESVANRLQGKILEW---------EKQNGGQNYF- 415
                             | :||: |   ::       ||:::              :  | ||   :
Db        423 ATTSVDSNDVSFEYDGKTKASEAAGIQATIKLDTGKVVDLTAADIIVTNDDVNAGQYSYQ 482

Qy        416 --DKGYDSRYAAYLANNLKFLSELNKELEAERVIAITQQRWDNNIGELAGITKLGERIKS 473
                | |  ::  |   || :  ::                    :: ::||    :    : :
Db        483 LSDAG-KAKLQAATGNNYQLTAD-------------------DLAKVAGTITITPAVTT 521

Qy        474 GKAYADAFE-DGKKVEA---GSNITLDAKTG-IIDISNSNGKKTQALHFTSPLLTAGTES 528
                :    :|| |||    :     |    |:    || ::|:: ::     :  |:  :  || |
Db        522 VDSSDVSFEYDGKTKASEAKGIQATIKLDTGKVVDLTAAD------IIVTNDDVNAGQYS 575

Qy        529 RERLTNGKYSYINKLKFGRVKNWQVT--------------------------DGEASSK 561
                :    ||     ||:      |:|:|                            :  :  :|

Db        576 YQLSDAGK----AKLQAATGNNYQLTADDLAKVMGTITITPAAVTADSNDLSFEYDGKTK 631

Qy        562 LDFSKVIQ---RVAETEGTDEIGLIVNAKAGNDDIFVGQ--------GKMNIDGGDGHDR 610
                   :| ||    :: ||| | ::     :      |||:  ||          || :      |::
Db        632 ASEAKGIQAMVKLGETEKTVDL-TSADIVVANDDVNAGQYSYQLSDAGKAKLQAATGNNY 690

Qy        611 VFYSKDGGFGNITVDGTSATEAGSYTVNRKVARGDI------YHEVVKRQETKVGKRTET 664
                              :| ||   |  ||: |:             |           |      |  | |
Db        691 ----------QLTADGL-AKVAGTITITPATTADSNDVSFEYDGKTKASEAK------- 732

Qy        665 IQYRDYELRKVGYGYQSTDNLKSVEEVIGSQFNDVFKGSKFNDIFHSGEGDDLLDGGAGD 724
                           | |:|  |   :|: :      |:                                      :
Db        733 -------------GIQATIKLGEIEKTVDLSSADII--------------------VAN 758

Qy        725 DRLFGGKGNDRLSGDEGDDLLDGGSGD------DVLNGGAGNDVYIFRKG---DGNDTLY 775
              | :    ||      || | |.    |    :|           :||:   :|:  : |    |     | |   :
Db        759 DGVIVGKYTYSLS-DSGKSKLQAATGSNYQLTTEVLDKVSGS-ITITPAGAIATGKDAHF 816

Qy        776 DGTGNDKLAFA---------------DANISDIMIERTKEGIIVKR------------- 806
              :  |   | :  |                   |       :||::    ::|:    :
Db        817 EYDGKTKASEAKGIQAILTIDGTEKTVDLTAADIVV--AEDGVDAGKYSYRLSDAGKSKL 874

Qy        807 -----NDH---------SGSINIPRWYIT--SNLQNYQSNKTDHKIEQ-------LIGK 842
                   :||          :|:|  |       |  ||  ::: |         |           ::|:
Db        875 QREAGSDHQLTADDLAEVTGTITITPAIATADSNDVSFEYNGKTKASEAEGIQATVMLGE 934

Qy        843 DGSYITSDQIDKILQDKKDGTVITSQELKKLADENKSQKLSASD-----IASSLNKLVGS 897
                |  :       | :: :    ||  |     :    :|:|    |::    :|:            |    |:|: |:
Db        935 SGQVVALTSADVVVVN--DG-VDAGKYSYQLSDAGKAKLQAATGNNYQLTADDLDKVTGT 991

Qy        898 MALFGTANSVSSN 910
                : :         :|  ||
Db        992 ITITPATTTVDSN 1004


RESULT 38
US-09-117-447-2
; Sequence 2, Application US/09117447
; Patent No. US20020168728A1
; GENERAL INFORMATION:
;  APPLICANT:  LUBITZ, Werner
;  APPLICANT:  SLEYTR, Uwe
;  APPLICANT:  KUEN, Beatrix
;  APPLICANT:  TRUPPE, Michaela
;  APPLICANT:  HOWORKA, Stefan
;  APPLICANT:  RESCH, Stepanka
;  APPLICANT:  SCHROLL, Gerhard
;  APPLICANT:  SARA, Margit
;  TITLE OF INVENTION:  RECOMBINANT EXPRESSION OF S-LAYER PROTEINS
;  FILE REFERENCE:  100564-08013
;  CURRENT APPLICATION NUMBER: US/09/117,447
;  CURRENT FILING DATE:  1998-12-02
;  PRIOR APPLICATION NUMBER:  PCT/EP97/00432
;  PRIOR FILING DATE:  1997-01-31
;  PRIOR APPLICATION NUMBER:  DE/196 03 649.6
;  PRIOR FILING DATE:  1996-02-01
;  NUMBER OF SEQ ID NOS:  10
;  SOFTWARE:  PatentIn Ver. 2.1
; SEQ ID NO 2
;   LENGTH: 1228
;   TYPE: PRT
;   ORGANISM: Bacillus stearothermophilus
US-09-117-447-2

  Query Match              3.8%;  Score 175.5;  DB 10;  Length 1228;
  Best Local Similarity   19.4%;  Pred. No. 0.0014;
  Matches  210;  Conservative  154;  Mismatches  355;  Indels  363;  Gaps   53;

Qy         15 LNSTKSGLKNLYLA-IPKDYDPQKGGTLNDFIKAADELGIARLAEEPNHTETAKKSVDTV 73
               :|      |: ||| : |:|:                 ||    :  | ||:|                   :|
Db         86 VNKAGGAKKDAYLADLQKEYETY-------VFKANPKSGEARVA----------TYIDAY 128

Qy         74 NQFLSLTQTGIAISATKLEKFLQKHSTNKLAKGLDSVEN---------------IDRKLG 118
```

```
Db    129 NY-----------ATKLDEMRQELEAAVQAKDLEKAEQYYHKIPYEIKTRTVILDRVYG 176

Qy    119 ------------KASNVLSTLSSFLGTALAGIELDSLIKKGDAAPDALAKASIDLINEI 165
Db    177 KTTRDLLRSTFKAKAQELRDSLIYDITVAMKAREVQDAVKAGNLDK---AKAAVDQINQY 233

Qy    166 IGNLSQSTQTIEAFSSQLAKLGSTISQAKGFSNIGNKLQNLNFSKT-NLGLEIITGLLSG 224
Db    234 LPKVT-----DAFKTELTEVAKKALDADE-AALTPKVESVSAINTQNKAVELTAVPVNG 286

Qy    225 -----ISAGFALADKNASTGKKVAAGFELSNQVIGNVTKAISSYVLAQRVAAGLSTTGAV 279
Db    287 TLKLQLSAAANEDTVNVNTVR--------IYKVDGNIPFALNT------ADVSLSTDGKT 332

Qy    280 AALITSSIMLAISPLAFMNAAD-KFNHANALDEFAKQFRK----FGYDGDHLLAEYQRGV 334
Db    333 ITVDAST--------PFENNTEYKVVVKGIKDKNGKEFKEDAFTFKLRNDAVVTQV---F 381

Qy    335 GTIEASLTTISTALGA---------------------------VSAGVSAAAVGSAVG 365
Db    382 GTNVTNNTSVNLAAGTFDTDDTLTVVFDKLLAPETVNSSNVTITDVETGKRIPVIASTSG 441

Qy    366 TPIALLV--AGVTG-----------LISGILEASKQAMFESVAN---------RLQGKIL 403
Db    442 STITITLKEALVTGKQYKLAINNVKTLTGYNAEAYELVFTANASAPTVATAPTTLGGTTL 501

Qy    404 E--------WEKQNGGQN-----YFDKGYDSRYA-----AYLANNLKFL---------SE 436
Db    502 STGSLTTNVWGKLAGGVNEAGTYYPGLQFTTTFATKLDESTLADNFVLVEKESGTVVASE 561

Qy    437 LNKELEAERVIAITQ-QRWDNNIGELAGITKLGERIKSGKAYADAFEDGKKVEAGSNITL 495
Db    562 LKYNADAKMVTLVPKADLKENTIYQI--------KIKKG------LKSDKGIELG---TV 604

Qy    496 DAKT-----------GIIDISNSNG----KKTQALHFTSPL------LTAGTESRERLTN 534
Db    605 NEKTYEFKTQDLTAPTVISVTSKNGDAGLKVTEAQEFTVKFSENLNTFNATTVSGSTITY 664

Qy    535 GKYSYINKLKFGRVKNWQVTDGEASSKLDFSKVI----QRVAETEGTDEIGLIVNAKAGN 590
Db    665 GQVAVVKA-------------GANLSALTASDIIPASVEAVTGQDGTYKVKVAAN----- 706

Qy    591 DDIFVGQGKMNIDGGDGHDRVFYSKDGGFGNITVDGTSATEAGSYTVNRKVARGDIYHEV 650
Db    707 ----------QLERNQGYKLVVFGK----------GATAPVKDAANAN-TLATNYIYTFT 745

Qy    651 VKRQETKVGKRTETIQYRDYELRKVGYGYQSTDNLKSVEEVIGSQFNDVFKGSKFNDIFH 710
Db    746 TEGQDVTAPTVTKVFK---------------GDSLKDADAV--TTLTNVDAGQKFTIQF- 787

Qy    711 SGEGDDLLDGGAGDDRLFGGKGNDRLSGDEGDDLLDGGSGDDVLNGGAGNDVYIFRKGDG 770
Db    788 ----SEELKTSSGS--LVGGKVT-------VEKLTNNGWVD----AGTGTTVSVAPKTDA 830

Qy    771 ND-------TLYDGTGNDKLAFADANISDIMIERTKEGI------IVKRNDHSGSINIPR 817
Db    831 NGKVTAAVVTLTGLDNNDK----DAKLRLVVDKSSTDGIADVAGNVIKEKD--------- 877

Qy    818 WYITSNLQNYQS-NKTDHKIEQLIGKDG-----SYITSDQIDKILQDKKDGTVITSQEL- 870
Db    878 -----ILIRYNSWRHTVASVKAAADKDGQNASAAFPTSTAID-----------TTKSLL 920

Qy    871 -----KKLADENKSQKLSASDIASSLNKLVGSMALFGTANSVSSNALQPITQPTQGILAP 925
Db    921 VEFNETDLA-EVKPENIVVKDAAGN--------AVAGTVTALDGSTNKFVFTPSQELKAG 971

Qy    926 SV 927
Db    972 TV 973
```

RESULT 39

```
US-09-882-227-522
; Sequence 522, Application US/09882227
; Publication No. US20030158396A1
; GENERAL INFORMATION:
;  APPLICANT: Kleanthous, Harold
;  APPLICANT:  Al-Garawi, Amal
;  APPLICANT:  Miller, Charles
;  APPLICANT:  Tomb, Jean-Francois
;  APPLICANT:  Ooomen, Raymond P.
;  TITLE OF INVENTION: Identification of Polynucleotides
;  TITLE OF INVENTION:  Encoding No. US20030158396A1el Helicobacter Polypeptides in the
;  TITLE OF INVENTION:  Genome
;  FILE REFERENCE: 06132/047002
;  CURRENT APPLICATION NUMBER: US/09/882,227
;  CURRENT FILING DATE:  2001-06-15
;  PRIOR APPLICATION NUMBER: US 08/902,615
;  PRIOR FILING DATE: 1997-07-29
;  NUMBER OF SEQ ID NOS: 638
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 522
;   LENGTH: 2893
;   TYPE: PRT
;   ORGANISM: Helicobacter pylori
US-09-882-227-522

  Query Match             3.8%;  Score 175;  DB 12;  Length 2893;
  Best Local Similarity   18.2%;  Pred. No. 0.0056;
  Matches  199;  Conservative  149;  Mismatches  378;  Indels  366;  Gaps   48;

Qy      1 MSNINVIKSNI--------QAGLNS--TKSGLKNLYLAIPKDYDPQKGGTLNDFIKAADE 50
Db   1843 LNQANIVSSQTDGIFSMLGQEGINKVFNQAGLANILGEVAVQSINKAGGLGNLIVNTLGS 1902

Qy     51 ---LGIARLAEEPNHT------------------ETAKKSVDTVNQFLSLTQTGI---- 84
Db   1903 NSVIGGYLTPEQKNQTLSQLLGQNNFDNLMNDSGLNTAIK--DLIRQKLGF-WTGLVGGL 1959

Qy     85 ----AISATKLEKFLQKHSTNKLAKGLDSVENIDRKLGKASNVLSTLSSFLGTALAGIEL 140
Db   1960 AGLGGIDLQNPEKLIGSMSINDLLS----------KKGLFNQITGFISANDIGQVISVML 2009

Qy    141 DSLIKKGDAAPDALAKASIDLINEIIG--------NLSQSTQTIEAFSSQLAK--LGSTI 190
Db   2010 QDIVKPSNALKNDVAALGKQMIGEFLGQDTLNSLESLLQNQQIKSVLDKVLAAKGLGPIY 2069

Qy    191 SQAKG--FSNIGNK-------------------------LQNLNFSKTNLG-------- 214
Db   2070 EQGLGDLIPNLGKKGLFAPYGLSQVWQKGDFSFNAQGNVFVQNSTFSNANGGTLSFNAGN 2129

Qy    215 ------------------LEIITGLLSGISAGFALADKNASTGKKVAAGFELSNQVIGNV 256
Db   2130 SLIFAGNNHIAFTNHAGTLQLLSDQVSNIN----ITTLNASNGLKINAANNNVSVSQGNL 2185

Qy    257 TKAISSYVLAQRVAAGLSTTGAVAALITSS-----------IMLAISPLAFMNAADKFN- 304
Db   2186 FVSASCAQQSDPTTANIANPCALSAQSTNGASSNNASNNAPIALSNNDESLMVAANDFNF 2245

Qy    305 ----HANALDEFAK-----------QFRKFGYDGDHLLAEYQ------------------ 331
Db   2246 SGNIYANGVVDFSKIKGSANIKNLYLYNNAQFQANNLTISNQAVLEKNASFVTNNLNIQG 2305

Qy    332 ----RGVGTIEASLTTISTALGAVSAGVSAAAVGSAVG-------------TPIALLVA- 373
Db   2306 AFNNNATQKIEVLQNLVIASNASLSTGIYGLEVGGALNNSGAIHFNLENTQTPTPLIQAE 2365

Qy    374 GVTGLIS------------------GILEASKQAMFESVANRLQGKILEW---------- 405
Db   2366 GIINLNTTQTPFMNVNNSMANNTTYTLLKSSRYIDYNINPNSLQSYLNLYTLININGNHI 2425

Qy    406 EKQNGGQNYF-------DKGY-------DSRYAAYLANNLKFLSELNKELE--------- 442
Db   2426 EEKNGALTYLGQRVLLQDKGLLLSVALPNSNNAS--QNNILSLSVLYNQVKMSCGDKAMD 2483
```

```
Qy        443 ------AERVIAITQQRWDNNIGELAG-----ITKLGERIKSGKAYADAFEDGKKVEAGS 491
Db       2484 FTPPTLQDYIVGIQGQSALNQIEAVGGNAIKWLSTLMMETKENPFFAPIYLKNHSLNEIL 2543

Qy        492 NITLDAKTGIIDISNSN--GKKTQALHFTSPLLTAGTESRERLTNGKYSYINKLKFGRVK 549
Db       2544 GVTKDLQNTASLISNPNFRDNATNLLE-----LASYTQQTSRLT------------KLS 2585

Qy        550 NWQVTDGEASSKLDFSKVIQRVAETEGTDEIGLIVNAKAGNDDIFVGQGKMN-------- 601
Db       2586 DFRSREGES----DFS--LLELKNKRFSDP---------NPEVFVKYSQLSKHPNNLWV 2629

Qy        602 --------IDGGDGHDRVFYSKDGGFG----NITVDGTSATEAGSYTVNRKVARGDIYHE 649
Db       2630 QGVGGASFISGGNG---TLYGLNAGYDRLVKNVILGGYVAYGYSDF-------NGNIMHS 2679

Qy        650 V------------VKR-------QETKVGKRT-------------ETIQYRDYELRKVG 676
Db       2680 LGNNVDVGMYARAFLKRNEFTLSANETYGGNATSINSSNSLLSVLNQRYNYNTWTTSVNG 2739

Qy        677 -YGYQSTDNLKSV--EEVIGSQFNDV-FKGSKFND------IFHSGEGDD---LLDGGAG 723
Db       2740 NYGYDFMFKQKSVVLKPQVGLSYHFIGLSGMKGNDAAYKQFLMHSNPSNESVLTLNMGLE 2799

Qy        724 DDRLFGGKGNDRLSGDEGDDLLDGGSGDDVLNGGAGNDVYIFRKGDGNDTLYDG-TGND- 781
Db       2800 SRKYFGKNSYYFVTARLGRDLLIKSKGSNTVR-FVGENTLLYRKGEVFNTFASVITGGEM 2858

Qy        782 ---KLAFADANI 790
Db       2859 HLWRLVYVNAGV 2870
```

RESULT 40
US-10-193-764-63
; Sequence 63, Application US/10193764
; Publication No. US20030133943A1
; GENERAL INFORMATION:
; APPLICANT: Loosmore, Sheena M.
; APPLICANT: Yang, Yan-Ping
; APPLICANT: Klein, Michel H.
; TITLE OF INVENTION: PROTECTIVE RECOMBINANT HAEMOPHILUS INFLUENZAE HIGH
; TITLE OF INVENTION: MOLECULAR WEIGHT PROTEINS
; FILE REFERENCE: 1038-1239MIS
; CURRENT APPLICATION NUMBER: US/10/193,764
; CURRENT FILING DATE: 2002-07-12
; PRIOR APPLICATION NUMBER: 09/167,568
; PRIOR FILING DATE: 1998-10-07
; NUMBER OF SEQ ID NOS: 91
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 63
;   LENGTH: 1536
;   TYPE: PRT
;   ORGANISM: Haemophilus influenzae
US-10-193-764-63

```
  Query Match             3.8%;  Score 174.5;  DB 12;  Length 1536;
  Best Local Similarity   20.8%;  Pred. No. 0.0023;
  Matches  138;  Conservative   80;  Mismatches  243;  Indels  203;  Gaps   28;

Qy        370 LLVAGVTGLISGILEASKQAM------------FESVANRLQGKILEWEKQNGGQNYFD 416
Db         54 LLSLGVTSIPQSVLASGLQGMDVVHGTATMQVDGNKTIIRNSVDDIINWKQFNIDQN--- 110

Qy        417 KGYDSRYAAYLANNLKFLSELNKELEAERVIAITQQRWDNNIGELAGITKLGERIKSGKA 476
Db        111 ------------EMVQFLQENNNSAVFNRVTS-------NQISQLKGILD-----SNGQV 146

Qy        477 Y---ADAFEDGKKVEAGSNITLDAKTGIIDISNSNGK-KTQALHFTSPLLTAGTESRERL 532
Db        147 FLINPNGITIGKDAIINTN-GFTAST--LDISNENIKARNFTFEQTKDKALAEIVNHGLI 203

Qy        533 TNGKYSYINKLKFGRVKNWQV--TDGEASSKLDFSKV-IQRVAETEGTDEIGLIVNAKAG 589
Db        204 TVGKDGSVN-LIGGKVKNEGVISVNGGSISLLAGQKITISDIINPTITYSIAAPENEAVN 262

Qy        590 NDDIFVGQGKMNIDGGDGHDRVFYSKDGGFGNITVDGTSATEAGSYTVNRKVARGDIYHE 649
Db        263 LGDIFAKGGNINVRAATIRNQ---------GKLSADSVSKDKSGNIVLSAKEGEAEIGGV 313

Qy        650 V-VKRQETKVGKRTETIQYRDYELRKVGYGYQSTDNLKSVEEVIGSQFNDVFKGSKFNDI 708
Db        314 ISAQNQQAKGGKLMIT-------------GDKVTLKTGAVIDLSGKEGGETYLG------ 354

Qy        709 FHSGEGDDLLDGGAGDDRLFGGKG------------NDRLSGDE--------GD-DLLD 746
Db        355 --------------GDERGEGKKGIQLAKKTSLEKGSTINVSGKEKGGRAIVWGDIALID 400

Qy        747 G-----GSGDDVLNGG----AGNDVYI------------------------FRKGDGND 772
Db        401 GNINAQGSGDIAKTGGFVETSGHDLFIKDNAIVDAKEWLLDPDNVSINAETAGRSNTSED 460

Qy        773 TLYDGTGN---------DKLAFADANISDIMIERTKEGIIVKRNDH-SGSINIPRWYIT- 821
Db        461 DEYTGSGNSASTPKRNKEKTTLTNTTLESILKKGTFVNITANQRIYVNSSINLSNGSLTL 520

Qy        822 -----------------------SNLQNYQSNKTD-HKIEQLIGKDGSYITSDQIDKIL 856
Db        521 WSEGRSGGGVEINNDITTGDDTRGANLTIYSGGWVDVHKNISLGAQGNINITAKQ--DIA 578

Qy        857 QDKKDGTVITSQELKKLADEN--KSQKLSASDIASSL------------NKLVGSMALF 901
Db        579 FEKGSNQVITGQGTITSGNQKGFRFNNVSLNGTGSGLQFTTKRTNKYAITNKFEGTLNIS 638

Qy        902 GTAN 905
Db        639 GKVN 642
```

Search completed: February 17, 2004, 10:18:33
Job time : 55 secs

```
OM protein - protein search, using sw model

Run on:         February 17, 2004, 10:09:49 ; Search time 46 Seconds
                                          (without alignments)
                                          3198.686 Million cell updates/sec

Title:          US-10-069-799-5
Perfect score:  4647
Sequence:       1 MSNINVIKSNIQAGLNSTKS..........SSNALQPITQPTQGILAPSV 927

Scoring table:  BLOSUM62
                Gapop 10.0 , Gapext 0.5

Searched:       1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters:      1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       A_Geneseq_19Jun03:*
                 1:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
                 2:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
                 3:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
                 4:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
                 5:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
                 6:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
                 7:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
                 8:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
                 9:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
                 10:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
                 11:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
                 12:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
                 13:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
                 14:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
                 15:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
                 16:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
                 17:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
                 18:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
                 19:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
                 20:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
                 21:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
                 22:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*
                 23:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2002.DAT:*
                 24:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2003.DAT:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	4647	100.0	927	22	AAB62110	M. bovis Dalton 2d
2	2334	50.2	1098	18	AAW13866	Chimeric protein #
3	2334	50.2	1098	21	AAB21073	Bovine IL-2/Pasteu
4	2332	50.2	1098	13	AAR22103	Bovine IL-2 - LKT
5	2329	50.1	953	14	AAR43865	Leukotoxin protein
6	2327	50.1	1098	15	AAR52747	Bovine IL-2/LKT ch
7	2326.5	50.1	924	14	AAR42385	Recombinant leukot
8	2326.5	50.1	924	14	AAR42380	Recombinant leukot
9	2326.5	50.1	924	14	AAR42378	Recombinant leukot
10	2326.5	50.1	926	12	AAR14482	LKT352. Pasteurel
11	2326.5	50.1	926	14	AAR34545	Leukotoxin 352 pro
12	2326.5	50.1	926	15	AAR50291	Recombinant leukot
13	2326.5	50.1	926	19	AAW79568	Leukotoxin 352 pol
14	2326.5	50.1	977	17	AAW03942	LKT-GnRH protein f
15	2326.5	50.1	977	19	AAW79569	LKT-GnRH chimeric
16	2326.5	50.1	1069	15	AAR52748	Bovine IFNgamma/LK
17	2326.5	50.1	1069	18	AAW13867	Chimeric protein #
18	2326.5	50.1	1069	21	AAB21074	Bovine gamma-IFN/P
19	2325	50.0	953	11	AAR07167	105kD PTX protein
20	2325	50.0	953	15	AAR60072	PtxA protein of Pa
21	2325	50.0	953	22	AAE04638	Pasteurella haemol
22	2323.5	50.0	943	14	AAR34546	Somatostatin-leuko
23	2318.5	49.9	936	14	AAR34547	GnRH-leukotoxin ge
24	2317.5	49.9	924	12	AAR10889	Leukotoxin 352 enc
25	2317.5	49.9	926	17	AAW03945	P. haemolytica tru
26	2313	49.8	953	12	AAR15159	Leukotoxin from P.
27	2312.5	49.8	951	14	AAR34548	Rotavirus VP4-leuk
28	2297.5	49.4	934	17	AAW07637	P. suis leukotoxin
29	2236.5	48.1	956	12	AAR12561	APPA haemolysin an
30	2236.5	48.1	956	18	AAW22156	ApxIIC protein. A
31	2236.5	48.1	956	21	AAY51410	A. pleuropneumonia
32	1924	41.4	1049	18	AAW22159	ApxIIIB protein.
33	1924	41.4	1049	21	AAY51412	A. pleuropneumonia
34	1924	41.4	1244	15	AAR54781	Leukotoxin AppIIIA
35	1763.5	37.9	1022	18	AAW22152	ApxIA protein. Ac
36	1763.5	37.9	1022	21	AAY51406	A. pleuropneumonia
37	1727	37.2	1023	16	AAR76991	LhaA (low homology
38	1464.5	31.5	608	22	AAE04636	Pasteurella haemol
39	1376.5	29.6	758	17	AAR86998	Enterohaemorrhagic
40	1159.5	25.0	450	22	AAE04637	Pasteurella haemol
41	904.5	19.5	544	19	AAW79570	LKT-GnRH chimeric
42	901	19.4	695	19	AAW79573	LKT-GnRH chimeric
43	901	19.4	695	21	AAY58361	Leukotoxin/gonadot
44	901	19.4	695	21	AAY58133	Gonadotropin relea
45	897.5	19.3	490	20	AAY33929	Leukotoxin carrier

ALIGNMENTS

```
RESULT 1
AAB62110
ID   AAB62110 standard; Protein; 927 AA.
XX
AC   AAB62110;
XX
DT   29-MAY-2001  (first entry)
XX
DE   M. bovis Dalton 2d RTX toxin A subunit.
XX
KW   Moraxella; antigen; immune response; infection; RTX toxin; vaccine;
KW   antibacterial; A subunit.
XX
OS   Moraxella bovis.
XX
PN   WO200116172-A1.
XX
PD   08-MAR-2001.
XX
PF   31-AUG-2000; 2000WO-AU01048.
XX
PR   31-AUG-1999;   99AU-0002571.
XX
PA   (CSIR ) COMMONWEALTH SCI & IND RES ORG.
PA   (UYME ) UNIV MELBOURNE.
XX
PI   Farn J,  Strugnell R,  Tennent J;
XX
DR   WPI; 2001-235092/24.
DR   N-PSDB; AAF57290.
XX
PT   Novel Moraxella bovis antigen useful in compositions for raising immune
```

```
PT   response in an animal, has protease, lipase or hemolysin activity  -
XX
PS   Claim 26; Fig 5; 60pp; English.
XX
CC   The invention relates to new Moraxella bovis antigens and nucleic acid
CC   sequences encoding these antigenic polypeptides. The antigenic
CC   polypeptides and polynucleotides are useful for raising an immune
CC   response in an animal directed against Moraxella, preferably against
CC   M. bovis or M. catarrhalis, and for treating Moraxella infections. The
CC   present sequence represents the amino acid sequence of the A subunit of
CC   the RTX toxin from M. bovis Dalton 2d.
XX
SQ   Sequence   927 AA;

  Query Match            100.0%;  Score 4647;  DB 22;  Length 927;
  Best Local Similarity  100.0%;  Pred. No. 4.1e-281;
  Matches  927;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MSNINVIKSNIQAGLNSTKSGLKNLYLAIPKDYDPQKGGTLNDFIKAADELGIARLAEEP 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MSNINVIKSNIQAGLNSTKSGLKNLYLAIPKDYDPQKGGTLNDFIKAADELGIARLAEEP 60

Qy         61 NHTETAKKSVDTVNQFLSLTQTGIAISATKLEKFLQKHSTNKLAKGLDSVENIDRKLGKA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 NHTETAKKSVDTVNQFLSLTQTGIAISATKLEKFLQKHSTNKLAKGLDSVENIDRKLGKA 120

Qy        121 SNVLSTLSSFLGTALAGIELDSLIKKGDAAPDALAKASIDLINEIIGNLSQSTQTIEAFS 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 SNVLSTLSSFLGTALAGIELDSLIKKGDAAPDALAKASIDLINEIIGNLSQSTQTIEAFS 180

Qy        181 SQLAKLGSTISQAKGFSNIGNKLQNLNFSKTNLGLEIITGLLSGISAGFALADKNASTGK 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 SQLAKLGSTISQAKGFSNIGNKLQNLNFSKTNLGLEIITGLLSGISAGFALADKNASTGK 240

Qy        241 KVAAGFELSNQVIGNVTKAISSYVLAQRVAAGLSTTGAVAALITSSIMLAISPLAFMNAA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 KVAAGFELSNQVIGNVTKAISSYVLAQRVAAGLSTTGAVAALITSSIMLAISPLAFMNAA 300

Qy        301 DKFNHANALDEFAKQFRKFGYDGDHLLAEYQRGVGTIEASLTTISTALGAVSAGVSAAAV 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 DKFNHANALDEFAKQFRKFGYDGDHLLAEYQRGVGTIEASLTTISTALGAVSAGVSAAAV 360

Qy        361 GSAVGTPIALLVAGVTGLISGILEASKQAMFESVANRLQGKILEWEKQNGGQNYFDKGYD 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GSAVGTPIALLVAGVTGLISGILEASKQAMFESVANRLQGKILEWEKQNGGQNYFDKGYD 420

Qy        421 SRYAAYLANNLKFLSELNKELEAERVIAITQQRWDNNIGELAGITKLGERIKSGKAYADA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 SRYAAYLANNLKFLSELNKELEAERVIAITQQRWDNNIGELAGITKLGERIKSGKAYADA 480

Qy        481 FEDGKKVEAGSNITLDAKTGIIDISNSNGKKTQALHFTSPLLTAGTESRERLTNGKYSYI 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 FEDGKKVEAGSNITLDAKTGIIDISNSNGKKTQALHFTSPLLTAGTESRERLTNGKYSYI 540

Qy        541 NKLKFGRVKNWQVTDGEASSKLDFSKVIQRVAETEGTDEIGLIVNAKAGNDDIFVGQGKM 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 NKLKFGRVKNWQVTDGEASSKLDFSKVIQRVAETEGTDEIGLIVNAKAGNDDIFVGQGKM 600

Qy        601 NIDGGDGHDRVFYSKDGGFGNITVDGTSATEAGSYTVNRKVARGDIYHEVVKRQETKVGK 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 NIDGGDGHDRVFYSKDGGFGNITVDGTSATEAGSYTVNRKVARGDIYHEVVKRQETKVGK 660

Qy        661 RTETIQYRDYELRKVGYGYQSTDNLKSVEEVIGSQFNDVFKGSKFNDIFHSGEGDDLLDG 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 RTETIQYRDYELRKVGYGYQSTDNLKSVEEVIGSQFNDVFKGSKFNDIFHSGEGDDLLDG 720

Qy        721 GAGDDRLFGGKGNDRLSGDEGDDLLDGGSGDDVLNGGAGNDVYIFRKGDGNDTLYDGTGN 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GAGDDRLFGGKGNDRLSGDEGDDLLDGGSGDDVLNGGAGNDVYIFRKGDGNDTLYDGTGN 780

Qy        781 DKLAFADANISDIMIERTKEGIIVKRNDHSGSININIPRWYI[illegible]SNLQNYQSNKTDHKIEQLI 840
              ||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db        781 DKLAFADANISDIMIERTKEGIIVKRNDHSGSINIPRWYI[illegible]SNLQNYQSNKTDHKIEQLI 840

Qy        841 GKDGSYITSDQIDKILQDKKDGTVITSQELKKLADENKSQKLSASDIASSLNKLVGSMAL 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GKDGSYITSDQIDKILQDKKDGTVITSQELKKLADENKSQKLSASDIASSLNKLVGSMAL 900

Qy        901 FGTANSVSSNALQPITQPTQGILAPSV 927
              |||||||||||||||||||||||||||
Db        901 FGTANSVSSNALQPITQPTQGILAPSV 927


RESULT 2
AAW13866
ID   AAW13866 standard; Protein; 1098 AA.
XX
AC   AAW13866;
XX
DT   25-MAR-2003  (updated)
DT   12-MAY-1997  (first entry)
XX
DE   Chimeric protein #1.
XX
KW   RTX cytotoxin; cytokine; immunogen; chimeric protein; cytokine; vaccine;
KW   interleukin-2; IL-2; gamma interferon; gamma IFN; leukotoxin; pneumonia;
KW   Pasteurella haemolytica; LKT352; respiratory disease; shipping fever;
KW   fibrinous pneumonia; cattle; therapy.
XX
OS   Synthetic.
XX
PN   US5594107-A.
XX
PD   14-JAN-1997.
XX
PF   20-DEC-1993;   93US-0170126.
XX
PR   20-DEC-1993;   93US-0170126.
PR   22-AUG-1990;   90US-0571301.
PR   16-OCT-1991;   91US-0777715.
XX
PA   (CIBA ) CIBA GEIGY CANADA LTD.
PA   (UYSA-) UNIV SASKATCHEWAN.
XX
PI   Campos M,  Hughes HPA,  Potter A;
XX
DR   WPI; 1997-099529/09.
DR   N-PSDB; AAT60032.
XX
PT   Immunogenic chimeric proteins comprising cytokine linked to RTX
PT   toxin - useful in vaccines, esp. against shipping fever in cattle
XX
PS   Claim 10; Column 25-32; 56pp; English.
XX
CC   AAW13866 and AAW13867 represent immunogenic chimeric proteins of the
CC   invention.  This sequence represents a chimeric protein containing the
CC   bovine interleukin-2 (IL-2) sequence and a leukotoxin sequence.  The
CC   chimeric proteins of the invention comprise a cytokine, selected from
CC   IL-2 and gamma interferon (gamma IFN), linked to at least one RTX toxin
CC   epitope (preferably the sequence shown in AAW13865).  The RTX toxin used
CC   to provide the epitope sequence is preferably a leukotoxin, especially
CC   the full-length Pasteurella haemolytica leukotoxin.  Alternatively, the
CC   leukotoxin is a truncated leukotoxin lacking leukotoxic activity,
CC   especially LKT352.  The chimeric proteins can be used for the production
CC   if vaccines against respiratory diseases such as pneumonia, particularly
CC   fibrinous pneumonia caused by P.haemolytica, including shipping fever in
CC   cattle.
CC   (Updated on 25-MAR-2003 to correct PF field.)
XX
SQ   Sequence   1098 AA;

  Query Match             50.2%;  Score 2334;  DB 18;  Length 1098;
  Best Local Similarity   50.1%;  Pred. No. 7.4e-137;
  Matches  465;  Conservative  175;  Mismatches  263;  Indels   26;  Gaps   13;

Qy          8 KSNIQAGLNSTKSGLKNLYLAIPKD--YDPQKGGTLNDFIKAADELGIARLAEEPNHTET 65
```

```
Db     173 QSLTQAG-SSLKTGAKKIILYIPQNYQYDTEQGNGLQDLVKAAEELGIEVQREERNNIAT 231

Qy      66 AKKSVDTVNQFLSLTQTGIAISATKLEKFLQKHSTNKLAKGLDSVENIDRKLGKASNVLS 125

Db     232 AQTSLGTIQTAIGLTERGIVLSAPQIDKLLQK---TKAGQALGSAESIVQNANKAKTVLS 288

Qy     126 TLSSFLGTALAGIELDSLIKKGDAAPDALAKASIDLINEIIGNLSQSTQTIEAFSSQLAK 185

Db     289 GIQSILGSVLAGMDLDEAL-QNNSNQHALAKAGLELTNSLIENIANSVKTLDEFGEQISQ 347

Qy     186 LGSTISQAKGFSNIGNKLQNL-NFSKTNLGLEIITGLLSGISAGFALADKNASTGKKVAA 244

Db     348 FGSKLQNIKGLGTLGDKLKNIGGLDKAGLGLDVISGLLSGATAALVLADKNASTAKKVGA 407

Qy     245 GFELSNQVIGNVTKAISSYVLAQRVAAGLSTTGAVAALITSSIMLAISPLAFMNAADKFN 304

Db     408 GFELANQVVGNITKAVSSYILAQRVAAGLSSTGPVAALIASTVSLAISPLAFAGIADKFN 467

Qy     305 HANALDEFAKQFRKFGYDGDHLLAEYQRGVGTIEASLTTISTALGAVSAGVSAAAVGSAV 364

Db     468 HAKSLESYAERFKKLGYDGDNLLAEYQRGTGTIDASVTAINTALAAIAGGVSAAAAGSVI 527

Qy     365 GTPIALLVAGVTGLISGILEASKQAMFESVANRLQGKILEWEKQNGGQNYFDKGYDSRYA 424

Db     528 ASPIALLVSGITGVISTILQYSKQAMFEHVANKIHNKIVEWEKNNHGKNYFENGYDARYL 587

Qy     425 AYLANNLKFLSELNKELEAERVIAITQQRWDNNIGELAGITKLGERIKSGKAYADAFEDG 484

Db     588 ANLQDNMKFLLNLNKELQAERVIAITQQQWDNNIGDLAGISRLGEKVLSGKAYVDAFEEG 647

Qy     485 KKVEAGSNITLDAKTGIIDISNSNGKKTQALHFTSPLLTAGTESRERLTNGKYSYINKLK 544

Db     648 KHIKADKLVQLDSANGIIDVSNSGKAKTQHILFRTPLLTPGTEHRERVQTGKYEYITKLN 707

Qy     545 FGRVKNWQVTDGEASSKLDFSKVIQRV-------AETEGTDEIGLIVNAKAGNDDIFVGQ 597

Db     708 INRVDSWKITDGAASSTFDLTNVVQRIGIELDNAGNVTKTKETKIIAKLGEGDDNVFVGS 767

Qy     598 GKMNIDGGDGHDRVFYSKDGGFGNITVDGTSATEAGSYTVNRKVARGDIYHEVVKRQETK 657

Db     768 GTTEIDGGEGYDRVHYSR-GNYGALTIDATKETEQGSYTVNRFVETGKALHEVTSTHTAL 826

Qy     658 VGKRTETIQYRDYELRKVGYGYQSTDNLKSVEEVIGSQFNDVFKGSKFNDIFHSGEGDDL 717

Db     827 VGNREEKIEYR-HSNNQHHAGYYTKDTLKAVEEIIGTSHNDIFKGSKFNDAFNGGDGVDT 885

Qy     718 LDGGAGDDRLFGGKGNDRLSGDEGDDLLDGGSGDDVLNGGAGNDVYIFRKGDGNDTLYDG 777

Db     886 IDGNDGNDRLFGGKGDDILDGGNGDDFIDGGKGNDLLHGGKGDDIFVHRKGDGNDIITDS 945

Qy     778 TGNDKLAFADANISDIMIERTKEGIIVKRNDHSGSINIPRWY----ITSNLQNYQSNKTD 833

Db     946 DGNDKLSFSDSNLKDLTFEKVKHNLVI-TNSKKEKVTIQNWFREADFAKEVPNYKATK-D 1003

Qy     834 HKIEQLIGKDGSYITSDQIDKILQDKKDGTVITSQELKKLADENKSQKLSASDIASSLNK 893

Db    1004 EKIEEIIGQNGERITSKQVDDLI--AKGNGKITQDELSKVVDNYELLKHS-KNVTNSLDK 1060

Qy     894 LVGSMALFGTANSVSSNALQPITQPTQGI 922

Db    1061 LISSVSAFTSSNDSRNVLVAPTSMLDQSL 1089
```

RESULT 3
AAB21073
ID   AAB21073 standard; Protein; 1098 AA.
XX
AC   AAB21073;
XX
DT   19-DEC-2000 (first entry)
XX
DE   Bovine IL-2/Pasteurella haemolytica leukotoxin fusion protein.
XX
KW   Bovine IL-2; interleukin-2; leukotoxin; LKT; respiratory disease;
KW   pneumonia; shipping fever; cattle; livestock; anti-Pasteurella vaccine;
KW   immunogen.
XX
OS   Chimeric - Bos taurus.
OS   Chimeric - Pasteurella haemolytica.
XX
PN   US6096320-A.
XX
PD   01-AUG-2000.
XX
PF   20-OCT-1997;   97US-0954418.
XX
PR   20-DEC-1993;   93US-0170126.
PR   22-JUL-1996;   96US-0681479.
PR   22-AUG-1990;   90US-0571301.
PR   16-OCT-1991;   91US-0777715.
XX
PA   (UYSA-) UNIV SASKATCHEWAN.
PA   (CIBA) CIBA GEIGY CANADA LTD.
XX
PI   Campos M, Hughes HPA, Potter A;
XX
DR   WPI; 2000-531543/48.
DR   N-PSDB; AAA72483.
XX
PT   Vaccine for stimulating immunity against pneumonia comprises chimeric
PT   protein comprising gamma-interferon and leukotoxin derived from
PT   Pasteurella haemolytica -
XX
PS   Example 1; Column 31-38; 56pp; English.
XX
CC   The invention relates to a novel vaccine composition comprising an
CC   immunogenic chimeric protein that comprises gamma-interferon (gamma-IFN)
CC   or an active fragment thereof, linked to an epitope of a Pasteurella
CC   haemolytica leukotoxin (LKT). Pasteurella species, especially Pasteurella
CC   haemolytica, are responsible for respiratory diseases in a range of
CC   agricultural animals, most particularly cattle, but also sheep, pigs,
CC   horses and fowl. Shipping fever is the most economically important
CC   respiratory disease associated with Pasteurella species, affecting
CC   15-30% of exposed cattle and resulting in a 2-5% mortality rate in the
CC   exposed population. The vaccine composition of the invention is
CC   is useful for preventing or ameliorating respiratory diseases such as
CC   pneumonia, particularly shipping fever pneumonia, in livestock. The
CC   present sequence represents a fusion protein comprising bovine
CC   interleukin-2 (IL-2) and Pasteurella haemolytica leukotoxin, which may
CC   also be used as an anti-Pasteurella vaccine.
XX
SQ   Sequence   1098 AA;

Query Match   50.2%; Score 2334; DB 21; Length 1098;
Best Local Similarity   50.1%; Pred. No. 7.4e-137;
Matches 465; Conservative 175; Mismatches 263; Indels   26; Gaps   13;

```
Qy       8 KSNIQAGLNSTKSGLKNLYLAIPKD--YDPQKGGTLNDFIKAADELGIARLAEEPNHTET 65

Db     173 QSLTQAG-SSLKTGAKKIILYIPQNYQYDTEQGNGLQDLVKAAEELGIEVQREERNNIAT 231

Qy      66 AKKSVDTVNQFLSLTQTGIAISATKLEKFLQKHSTNKLAKGLDSVENIDRKLGKASNVLS 125

Db     232 AQTSLGTIQTAIGLTERGIVLSAPQIDKLLQK---TKAGQALGSAESIVQNANKAKTVLS 288

Qy     126 TLSSFLGTALAGIELDSLIKKGDAAPDALAKASIDLINEIIGNLSQSTQTIEAFSSQLAK 185

Db     289 GIQSILGSVLAGMDLDEAL-QNNSNQHALAKAGLELTNSLIENIANSVKTLDEFGEQISQ 347

Qy     186 LGSTISQAKGFSNIGNKLQNL-NFSKTNLGLEIITGLLSGISAGFALADKNASTGKKVAA 244

Db     348 FGSKLQNIKGLGTLGDKLKNIGGLDKAGLGLDVISGLLSGATAALVLADKNASTAKKVGA 407

Qy     245 GFELSNQVIGNVTKAISSYVLAQRVAAGLSTTGAVAALITSSIMLAISPLAFMNAADKFN 304
```

```
Db        408 GFELANQVVGNITKAVSSYILAQRVAAGLSSTGPVAALIASTVSLAISPLAFAGIADKFN 467

Qy        305 HANALDEFAKQFRKFGYDGDHLLAEYQRGVGTIEASLTTISTALGAVSAGVSAAAVGSAV 364
Db        468 HAKSLESYAERFKKLGYDGDNLLAEYQRGTGTIDASVTAINTALAAIAGGVSAAAAGSVI 527

Qy        365 GTPIALLVAGVTGLISGILEASKQAMFESVANRLQGKILEWEKQNGGQNYFDKGYDSRYA 424
Db        528 ASPIALLVSGITGVISTILQYSKQAMFEHVANKIHNKIVEWEKNNHGKNYFENGYDARYL 587

Qy        425 AYLANNLKFLSELNKELEAERVIAITQQRWDNNIGELAGITKLGERIKSGKAYADAFEDG 484
Db        588 ANLQDNMKFLLNLNKELQAERVIAITQQQWDNNIGDLAGISRLGEKVLSGKAYVDAFEEG 647

Qy        485 KKVEAGSNITLDAKTGIIDISNSNGKKTQALHFTSPLLTAGTESRERLTNGKYSYINKLK 544
Db        648 KHIKADKLVQLDSANGIIDVSNSGKAKTQHILFRTPLLTPGTEHRERVQTGKYEYITKLN 707

Qy        545 FGRVKNWQVTDGEASSKLDFSKVIQRV-------AETEGTDEIGLIVNAKAGNDDIFVGQ 597
Db        708 INRVDSWKITDGAASSTFDLTNVVQRIGIELDNAGNVTKTKETKIIAKLGEGDDNVFVGS 767

Qy        598 GKMNIDGGDGHDRVFYSKDGGFGNITVDGTSATEAGSYTVNRKVARGDIYHEVVKRQETK 657
Db        768 GTTEIDGGEGYDRVHYSR-GNYGALTIDATKETEQGSYTVNRFVETGKALHEVTSTHTAL 826

Qy        658 VGKRTETIQYRDYELRKVGYGYQSTDNLKSVEEVIGSQFNDVFKGSKFNDIFHSGEGDDL 717
Db        827 VGNREEKIEYR-HSNNQHHAGYYTKDTLKAVEEIIGTSHNDIFKGSKFNDAFNGGDGVDT 885

Qy        718 LDGGAGDDRLFGGKGNDRLSGDEGDDLLDGGSGDDVLNGGAGNDVYIFRKGDGNDTLYDG 777
Db        886 IDGNDGNDRLFGGKGDDILDGGNGDDFIDGGKGNDLLHGGKGDDIFVHRKGDGNDIITDS 945

Qy        778 TGNDKLAFADANISDIMIERTKEGIIVKRNDHSGSINIPRWY----ITSNLQNYQSNKTD 833
Db        946 DGNDKLSFSDSNLKDLTFEKVKHNLVI-TNSKKEKVTIQNWFREADFAKEVPNYKATK-D 1003

Qy        834 HKIEQLIGKDGSYITSDQIDKILQDKKDGTVITSQELKKLADENKSQKLSASDIASSLNK 893
Db       1004 EKIEEIIGQNGERITSKQVDDLI--AKGNGKITQDELSKVVDNYELLKHS-KNVTNSLDK 1060

Qy        894 LVGSMALFGTANSVSSNALQPITQPTQGI 922
Db       1061 LISSVSAFTSSNDSRNVLVAPTSMLDQSL 1089
```

RESULT 4
AAR22103
ID   AAR22103 standard; Protein; 1098 AA.
XX
AC   AAR22103;
XX
DT   06-JUL-1992 (first entry)
XX
DE   Bovine IL-2 - LKT fusion protein.
XX
KW   Interleukin 2; leuko-toxin; vaccine; pneumonia; respiratory
KW   diseases.
XX
OS   Pasteurella haemolytica.
OS   Bos taurus.
XX
PN   WO9203558-A.
XX
PD   05-MAR-1992.
XX
PF   22-AUG-1991;   91WO-CA00299.
XX
PR   22-AUG-1990;   90US-0571301.
XX
PA   (POTT/) POTTER A.
XX
PI   Potter A, Campos M, Hughes HPA;
XX
DR   WPI; 1992-096901/12.
DR   N-PSDB; AAQ22771.
XX
PT   Interleukin 2-leuko-toxin gene fusion - encodes fusion protein
PT   useful as vaccine for animal pneumonia
XX
PS   Claim 20; Fig 3; 68pp; English.
XX
CC   The IL-2-LKT protein was encoded by a chimeric gene contg. the
CC   bovine IL-2 gene fused to the DNA encoding at least one epitope of
CC   leukotoxin from P. haemolytica. IL-2-LKT was gel purified and
CC   ligated into the expression vector pGH433 lacI. The resulting
CC   clone pAA356 (ATCC 68386) contd. the desired gene fusion under the
CC   control of the E. coli lac promoter. The protein produced by the
CC   gene fusion is useful in a vaccine compsn. with a pharmaceutically
CC   acceptable vehicle, e.g. a carrier homologous to a rotavirus VP6
CC   inner capsid protein. The vaccine can be used for preventing or
CC   ameliorating respiratory diseases in animals e.g. shipping fever
CC   or pneumonia.
CC   See also AAR24124,5.
XX
SQ   Sequence   1098 AA;

```
  Query Match             50.2%;  Score 2332;  DB 13;  Length 1098;
  Best Local Similarity   50.1%;  Pred. No. 9.9e-137;
  Matches  465;  Conservative  174;  Mismatches  264;  Indels   26;  Gaps   13;

Qy          8 KSNIQAGLNSTKSGLKNLYLAIPKD--YDPQKGGTLNDFIKAADELGIARLAEEPNHTET 65
Db        173 QSLTQAG-SSLKTGAKKIILYIPQNYQYDTEQGNGLQDLVKAAEELGIEVQREERNNIAT 231

Qy         66 AKKSVDTVNQFLSLTQTGIAISATKLEKFLQKHSTNKLAKGLDSVENIDRKLGKASNVLS 125
Db        232 AQTSLGTIQTAIGLTERGIVLSAPQIDKLLQK---TKAGQALGSAESIVQNANKAKTVLS 288

Qy        126 TLSSFLGTALAGIELDSLIKKGDAAPDALAKASIDLINEIIGNLSQSTQTIEAFSSQLAK 185
Db        289 GIQSILGSVLAGMDLDEAL-QNNSNQHALAKAGLELTNSLIENIANSVKTLDEFGEQISQ 347

Qy        186 LGSTISQAKGFSNIGNKLQNL-NFSKTNLGLEIITGLLSGISAGFALADKNASTGKKVAA 244
Db        348 FGSKLQNIKGLGTLGDKLKNIGGLDKAGLGLDVISGLLSGATAALVLADKNASTAKKVGA 407

Qy        245 GFELSNQVIGNVTKAISSYVLAQRVAAGLSTTGAVAALITSSIMLAISPLAFMNAADKFN 304
Db        408 GFELANQVVGNITKAVSSYILAQRVAAGLSSTGPVAALIASTVSLAISPLAFAGIADKFN 467

Qy        305 HANALDEFAKQFRKFGYDGDHLLAEYQRGVGTIEASLTTISTALGAVSAGVSAAAVGSAV 364
Db        468 HAKSLESYAERFKKLGYDGDNLLAEYQRGTGTIDASVTAINTALAAIAGGVSAAAAGSVI 527

Qy        365 GTPIALLVAGVTGLISGILEASKQAMFESVANRLQGKILEWEKQNGGQNYFDKGYDSRYA 424
Db        528 ASPIALLVSGITGVISTILQYSKQAMFEHVANKIHNKIVEWEKNNHGKNYFENGYDARYL 587

Qy        425 AYLANNLKFLSELNKELEAERVIAITQQRWDNNIGELAGITKLGERIKSGKAYADAFEDG 484
Db        588 ANLQDNMKFLLNLNKELQAERVIAITQQQWDNNIGDLAGISRLGEKVLSGKAYVDAFEEG 647

Qy        485 KKVEAGSNITLDAKTGIIDISNSNGKKTQALHFTSPLLTAGTESRERLTNGKYSYINKLK 544
Db        648 KHIKADKLVQLDSANGIIDVSNSGKAKTQHILFRTPLLTPGTEHRERVQTGKYEYITKLN 707

Qy        545 FGRVKNWQVTDGEASSKLDFSKVIQRV-------AETEGTDEIGLIVNAKAGNDDIFVGQ 597
Db        708 INRVDSWKITDGAASSTFDLTNVVQRIGIELDNAGNVTKTKETKIIAKLGEGDDNVFVGS 767

Qy        598 GKMNIDGGDGHDRVFYSKDGGFGNITVDGTSATEAGSYTVNRKVARGDIYHEVVKRQETK 657
Db        768 GTTEIDGGEGYDRVHYSR-GNYGALTIDATKETEQGSYTVNRFVETGKALHEVTSTHTAL 826
```

```
Qy        658 VGKRTETIQYRDYELRKVGYGYQSTDNLKSVEEVIGSQFNDVFKGSKFNDIFHSGEGDDL 717
Db        827 VGNREEKIEYR-HSNNQHHAGYYTKDTLKAVEEIIGTSHNDIFKGSKFNDAFNGGDGVDT 885

Qy        718 LDGGAGDDRLFGGKGNDRLSGDEGDDLLDGGSGDDVLNGGAGNDVYIFRKGDGNDTLYDG 777
Db        886 IDGNDGNDRLFGGKGDDILDGGNGDDFIDGGKGNDLLHGGKGDDIFVHRKGDGNDIITDS 945

Qy        778 TGNDKLAFADANISDIMIERTKEGIIVKRNDHSGSINIPRWY----ITSNLQNYQSNKTD 833
Db        946 DGNDKLSFSDSNFKDLTFEKVKHNLVI-TNSKKEKVTIQNWFREADFAKEVPNYKATK-D 1003

Qy        834 HKIEQLIGKDGSYITSDQIDKILQDKKDGTVITSQELKKLADENKSQKLSASDIASSLNK 893
Db       1004 EKIEEIIGQNGERITSKQVDDLI--AKGNGKITQDELSKVVDNYELLKHS-KNVTNSLDK 1060

Qy        894 LVGSMALFGTANSVSSNALQPITQPTQGI 922
Db       1061 LISSVSAFTSSNDSRNVLVAPTSMLDQSL 1089
```

RESULT 5
AAR43865
ID   AAR43865 standard; Protein; 953 AA.
XX
AC   AAR43865;
XX
DT   25-MAR-2003 (updated)
DT   22-DEC-1993 (first entry)
XX
DE   Leukotoxin protein.
XX
KW   Antigenic determinants; peptides; vaccine; supernatant; leukotoxin.
XX
OS   Pasteurella haemolytica.
XX
PN   CA2081950-A.
XX
PD   02-MAY-1993.
XX
PF   02-NOV-1992;   92CA-2081950.
XX
PR   01-NOV-1991;   91US-0786662.
XX
PA   (UYGU-) UNIV GUELPH.
XX
PI   Conlon JA, Lo RYC, Shewen PE, Strathdee CA;
XX
DR   WPI; 1993-227822/29.
DR   N-PSDB; AAQ46410.
XX
PT   Vaccine for preventing Pasteurella haemolytica infections -
PT   contains leukotoxin protein antigenic determinants and P.
PT   haemolytica culture supernatant
XX
PS   Claim 1; Figure 1; 40pp; English.
XX
CC   One or more biologically pure antigenic determinants of the
CC   leukotoxin protein comprising of at least six amino acids is a
CC   component of a new vaccine for animals. The vaccine elicits an
CC   enhanced immune response after challenge with Pasteurella
CC   haemolytica. The vaccines other component is a bacterial free
CC   culture supernatant derived from a culture of Pasteurella
CC   haemolytica.
CC   (Updated on 25-MAR-2003 to correct PF field.)
XX
SQ   Sequence   953 AA;

  Query Match   50.1%; Score 2329; DB 14; Length 953;
  Best Local Similarity   49.9%; Pred. No. 1.3e-136;
  Matches 464; Conservative 175; Mismatches 264; Indels   26; Gaps   13;

```
Qy          8 KSNIQAGLNSTKSGLKNLYLAIPKD--YDPQKGGTLNDFIKAADELGIARLAEEPNHTET 65
Db         28 QSLTQAG-SSLKTGAKKIILYIPQNYQYDTEQGNGLQDLVKAAEELGIEVQREERNNIAT 86

Qy         66 AKKSVDTVNQFLSLTQTGIAISATKLEKFLQKHSTNKLAKGLDSVENIDRKLGKASNVLS 125
Db         87 AQTSLGTIQTAIGLTERGIVLSAPQIDKLLQK---TKAGQALGSAESIVQNANKAKTVLS 143

Qy        126 TLSSFLGTALAGIELDSLIKKGDAAPDALAKASIDLINEIIGNLSQSTQTIEAFSSQLAK 185
Db        144 GIQSILGSVLAGMDLDEAL-QNNSNQHALAKAGLELTNSLIENIANSVKTLDEFGEQISQ 202

Qy        186 LGSTISQAKGFSNIGNKLQNL-NFSKTNLGLEIITGLLSGISAGFALADKNASTGKKVAA 244
Db        203 FGSKLQNIKGLGTLGDKLKNIGGLDKAGLGLDVISGLLSGATAALVLADKNASTAKKVGA 262

Qy        245 GFELSNQVIGNVTKAISSYVLAQRVAAGLSTTGAVAALITSSIMLAISPLAFMNAADKFN 304
Db        263 GFELANQVVGNITKAVSSYILAQRVAAGLSSTGPVAALIASTVSLAISPLAFAGIADKFN 322

Qy        305 HANALDEFAKQFRKFGYDGDHLLAEYQRGVGTIEASLTTISTALGAVSAGVSAAAVGSAV 364
Db        323 HAKSLESYAERFKKLGYDGDNLLAEYQRGTGTIDASVTAINTALAAIAGGVSAAAAGSVI 382

Qy        365 GTPIALLVAGVTGLISGILEASKQAMFESVANRLQGKILEWEKQNGGQNYFDKGYDSRYA 424
Db        383 ASPIALLVSGITGVISTILQYSKQAMFEHVANKIHNKIVEWEKNNHGKNYFENGYDARYL 442

Qy        425 AYLANNLKFLSELNKELEAERVIAITQQRWDNNIGELAGITKLGERIKSGKAYADAFEDG 484
Db        443 ANLQDNMKFLLNLNKELQAERVIAITQQQWDNNIGDLAGISRLGEKVLSGKAYVDPFEEG 502

Qy        485 KKVEAGSNITLDAKTGIIDISNSNGKKTQALHFTSPLLTAGTESRERLTNGKYSYINKLK 544
Db        503 KHIKADKLVQLDSANGIIDVSNSGKAKTQHILFRTPLLTPGTEHRERVQTGKYEYITKLN 562

Qy        545 FGRVKNWQVTDGEASSKLDFSKVIQRV-------AETEGTDEIGLIVNAKAGNDDIFVGQ 597
Db        563 INRVDSWKITDGAASSTFDLTNVVQRIGIELDNAGNVTKTKETKIIAKLGEGDDNVFVGS 622

Qy        598 GKMNIDGGDGHDRVFYSKDGGFGNITVDGTSATEAGSYTVNRKVARGDIYHEVVKRQETK 657
Db        623 GTTEIDGGEGYDRVHYSR-GNYGALTIDATKETEQGSYTVNRFVETGKALHEVTSTHTAL 681

Qy        658 VGKRTETIQYRDYELRKVGYGYQSTDNLKSVEEVIGSQFNDVFKGSKFNDIFHSGEGDDL 717
Db        682 VGNREEKIEYR-HSNNQHHAGYYTKDTLKAVEEIIGTSHNDIFKGSKFNDAFNGGDGVDT 740

Qy        718 LDGGAGDDRLFGGKGNDRLSGDEGDDLLDGGSGDDVLNGGAGNDVYIFRKGDGNDTLYDG 777
Db        741 IDGNDGNDRLFGGKGDDILDGGNGDDFIDGGKGNDLLHGGKGDDIFVHRKGDGNDIITDS 800

Qy        778 TGNDKLAFADANISDIMIERTKEGIIVKRNDHSGSINIPRWY----ITSNLQNYQSNKTD 833
Db        801 DGNDKLSFSDSNLKDLTFEKVKHNLVI-TNSKKEKVTIQNWFREADFAKEVPNYKATK-D 858

Qy        834 HKIEQLIGKDGSYITSDQIDKILQDKKDGTVITSQELKKLADENKSQKLSASDIASSLNK 893
Db        859 EKIEEIIGQNGERITSKQVDDLI--AKGNGKITQDELSKVVDNYELLKHS-KNVTNSLDK 915

Qy        894 LVGSMALFGTANSVSSNALQPITQPTQGI 922
Db        916 LISSVSAFTSSNDSRNVLVAPTSMLDQSL 944
```

RESULT 6
AAR52747
ID   AAR52747 standard; Protein; 1098 AA.
XX
AC   AAR52747;
XX
DT   25-MAR-2003 (updated)
DT   01-JUL-1994 (first entry)
XX

```
DE   Bovine IL-2/LKT chimeric protein encoded by plasmid pAA356.
XX
KW   Bovine; interleukin-2; IL2; P. haemolytica; leukotoxin; LTK;
KW   ltkA; chromosome walking; fusion protein; vaccine;
KW   monoclonal; polyclonal; antibody; ss.
XX
OS   Pasteurella heamolytica - chimera.
OS   Bos taurus - chimera.
XX
PN   US5273889-A.
XX
PD   28-DEC-1993.
XX
PF   16-OCT-1991;   91US-0777715.
XX
PR   22-AUG-1990;   90US-0571301.
PR   16-OCT-1991;   91US-0777715.
XX
PA   (CIBA ) CIBA GEIGY CANADA LTD.
PA   (UYSA-) UNIV SASKATCHEWAN.
XX
PI   Campos M,  Hughes HPA,  Potter A;
XX
DR   WPI; 1994-006687/01.
DR   N-PSDB; AAQ54212.
XX
PT   Immunogenic fusion proteins of gamma-interferon and immunogenic
PT   leukotoxin - used in vaccines and to raise monoclonal and polyclonal
PT   antibodies
XX
PS   Disclosure; Fig 3; 56pp; English.
XX
CC   This sequence represents a fusion between bovine interleukin-2 (IL2)
CC   and Pasteurella haemolytica leukotoxin (LTK).  The leukotoxin gene,
CC   ltkA, was isolated from a gene library of P. haemolytica by
CC   chromosome walking.  Immunogenic fusion proteins such as this can be
CC   used in vaccine compositions.  It can also be used to raise mono-
CC   and polyclonal antibodies.
CC   (Updated on 25-MAR-2003 to correct PF field.)
XX
SQ   Sequence   1098 AA;

  Query Match            50.1%;  Score 2327;  DB 15;  Length 1098;
  Best Local Similarity  49.9%;  Pred. No. 2e-136;
  Matches  464;  Conservative  175;  Mismatches  264;  Indels   26;  Gaps   13;

Qy          8 KSNIQAGLNSTKSGLKNLYLAIPKD--YDPQKGGTLNDFIKAADELGIARLAEEPNHTET 65
Db        173 QSLTQAG-SSLKTGAKKIILYIPQNYQYDTEQGNGLQDLVKAAEELGIEVQREERNNIAT 231

Qy         66 AKKSVDTVNQFLSLTQTGIAISATKLEKFLQKHSTNKLAKGLDSVENIDRKLGKASNVLS 125
Db        232 AQTSLGTIQTAIGLTERGIVLSAPQIDKLLQK---TKAGQALGSAESIVQNANKAKTVLS 288

Qy        126 TLSSFLGTALAGIELDSLIKKGDAAPDALAKASIDLINEIIGNLSQSTQTIEAFSSQLAK 185
Db        289 GIQSILGSVLAGMDLDEAL-QNNSNQHALAKAGLELTNSLIENIANSVKTLDEFGEQISQ 347

Qy        186 LGSTISQAKGFSNIGNKLQNL-NFSKTNLGLEIITGLLSGISAGFALADKNASTGKKVAA 244
Db        348 FGSKLQNIKGLGTLGDKLKNIGGLDKAGLGLDVISGLLSGATAALVLADKNASTAKKVGA 407

Qy        245 GFELSNQVIGNVTKAISSYVLAQRVAAGLSTTGAVAALITSSIMLAISPLAFMNAADKFN 304
Db        408 GFELANQVVGNITKAVSSYILAQRVAAGLSSTGPVAALIASTVSLAISPLAFAGIADKFN 467

Qy        305 HANALDEFAKQFRKFGYDGDHLLAEYQRGVGTIEASLTTISTALGAVSAGVSAAAVGSAV 364
Db        468 HAKSLESYAERFKKLGYDGDNLLAEYQRGTGTIDASVTAINTALAAIAGGVSAAAAGSVI 527

Qy        365 GTPIALLVAGVTGLISGILEASKQAMFESVANRLQGKILEWEKQNGGQNYFDKGYDSRYA 424
Db        528 ASPIALLVSGITGVISTILQYSKQAMFEHVANKIHNKIVEWEKNNHGKNYFENGYDARYL 587

Qy        425 AYLANNLKFLSELNKELEAERVIAITQQRWDNNIGELAGITKLGERIKSGKAYADAFEDG 484
Db        588 ANLQDNMKFLLNLNKELQAERVIAITQQQWDNNIGDLAGISRLGEKVLSGKAYVDAFEEG 647

Qy        485 KKVEAGSNITLDAKTGIIDISNSNGKKTQALHFTSPLLTAGTESRERLTNGKYSYINKLK 544
Db        648 KHIKADKLVQLDSANGIIDVSNSGKAKTQHILFRTPLLTPGTEHRERVQTGKYEYITKLN 707

Qy        545 FGRVKNWQVTDGEASSKLDFSKVIQRV-------AETEGTDEIGLIVNAKAGNDDIFVGQ 597
Db        708 INRVDSWKITDGAASSTFDLTNVVQRIGIELDNAGNVTKTKETKIIAKLGEGDDNVFVGS 767

Qy        598 GKMNIDGGDGHDRVFYSKDGGFGNITVDGTSATEAGSYTVNRKVARGDIYHEVVKRQETK 657
Db        768 GTTEIDGGEGYDRVHYSR-GNYGALTIDATKETEQGSYTVNRFVETGKALHEVTSTHTAL 826

Qy        658 VGKRTETIQYRDYELRKVGYGYQSTDNLKSVEEVIGSQFNDVFKGSKFNDIFHSGEGDDL 717
Db        827 VGNREEKIEYR-HSNNQHHAGYYTKDTLKAVEEIIGTSHNDIFKGSKFNDAFNGGDGVDT 885

Qy        718 LDGGAGDDRLFGGKGNDRLSGDEGDDLLDGGSGDDVLNGGAGNDVYIFRKGDGNDTLYDG 777
Db        886 IDGNDGNDPLFGGKGDDILDGGNGDDFIDGGKGNDLLHGGKGDDIFVHRKGDGNDIITDS 945

Qy        778 TGNDKLAFADANISDIMIERTKEGIIVKRNDHSGSINIPRWY----ITSNLQNYQSNKTD 833
Db        946 DGNDKLSFSDSNLKDLTFEKVKHNLVI-TNSKKEKVTIQNWFREADFAKEVPNYKATK-D 1003

Qy        834 HKIEQLIGKDGSYITSDQIDKILQDKKDGTVITSQELKKLADENKSQKLSASDIASSLNK 893
Db       1004 EKIEEIIGQNGERITSKQVDDLI--AKGNGKITQDELSKVVDNYELLKHS-KNVTNSLDK 1060

Qy        894 LVGSMALFGTANSVSSNALQPITQPTQGI 922
Db       1061 LISSVSAFTSSNDSRNVLVAPTSMLDQSL 1089

RESULT 7
AAR42385
ID   AAR42385 standard; Protein; 924 AA.
XX
AC   AAR42385;
XX
DT   25-MAR-2003  (updated)
DT   19-APR-1994  (first entry)
XX
DE   Recombinant leukotoxin peptide from plasmid pCRR28.
XX
KW   Haemophilus somnus; immunogenic; haemolysin; LppB; LppC;
KW   thromboembolic meningoencephalitis; septicaemia; arthritis;
KW   pneumonia; lktA gene; haemin-binding protein; fusion protein.
XX
OS   Pasteurella haemolytica.
XX
PN   WO9321323-A1.
XX
PD   28-OCT-1993.
XX
PF   05-APR-1993;   93WO-CA00135.
XX
PR   09-APR-1992;   92US-0865050.
PR   04-JUN-1992;   92US-0893424.
PR   04-JUN-1992;   92US-0893426.
PR   29-MAR-1993;   93US-0038287.
PR   29-MAR-1993;   93US-0038288.
PR   29-MAR-1993;   93US-0038719.
XX
PA   (UYSA-) UNIV SASKATCHEWAN.
XX
PI   Harland RJ,  Pfeiffer CG,  Pontarollo RA,  Potter AA;
PI    Rioux C,  Theisen M;
XX
```

```
DR   WPI; 1993-351733/44.
DR   N-PSDB; AAQ51086.
XX
PT   Haemophilus somnus immunogenic proteins used in vaccines -
PT   selected from haemin-binding protein, haemolysin, LppB and LppC,
PT   and corresp. DNA
XX
PS   Disclosure; Fig 11; 119pp; English.
XX
CC   The lppB gene protein was expressed in E. coli as a fusion to the
CC   Pasteurella haemolytica leukotoxin gene lktA coded for by plasmid
CC   pAA352.  The llpB gene fragment was taken from pMS11.  LppB can be
CC   used in vaccines for preventing or treating H. somnus infections,
CC   which cause thromboembolic meningo-encephalitis, septicaemia, arthritis
CC   and pneumonia in vertebrates.
CC   See also AAR42370-86.
CC   (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ   Sequence   924 AA;

  Query Match             50.1%;  Score 2326.5;  DB 14;  Length 924;
  Best Local Similarity   50.0%;  Pred. No. 1.7e-136;
  Matches  461;  Conservative  174;  Mismatches  262;  Indels   25;  Gaps   12;

Qy         15 LNSTKSGLKNLYLAIPKD--YDPQKGGTLNDFIKAADELGIARLAEEPNHTETAKKSVDT 72
              |:  |:| | : ||:: || ::| | | :|||:|||  || |: ||: |: |
Db          7 LSFPKTGAKKIILYIPQNYQYDTEQGNGLQDLVKAAEELGIEVQREERNNIATAQTSLGT 66

Qy         73 VNQFLSLTQTGIAISATKLEKFLQKHSTNKLAKGLDSVENIDRKLGKASNVLSTLSSFLG 132
              :   : ||: || :|| :::| |||    |  : | | |:| :   ||  ||| : | ||
Db         67 IQTAIGLTERGIVLSAPQIDKLLQK---TKAGQALGSAESIVQNANKAKTVLSGIQSILG 123

Qy        133 TALAGIELDSLIKKGDAAPDALAKASIDLINEIIGNLSQSTQTIEAFSSQLAKLGSTISQ 192
              : |||::||  : : ::  ||||| ::| | :| |:: | :|:: |  |::: || :
Db        124 SVLAGMDLDEAL-QNNSNQHALAKAGLELTNSLIENIANSVKTLDEFGEQISQFGSKLQN 182

Qy        193 AKGFSNIGNKLQNL-NFSKTNLGLEIITGLLSGISAGFALADKNASTGKKVAAGFELSNQ 251
               ||  :|:||:|:   |  |||::|:||||| :|  |||||||| ||| |||||:||
Db        183 IKGLGTLGDKLKNIGGLDKAGLGLDVISGLLSGATAALVLADKNASTAKKVGAGFELANQ 242

Qy        252 VIGNVTKAISSYVLAQRVAAGLSTTGAVAALITSSIMLAISPLAFMNAADKFNHANALDE 311
              |:||:|||:|||:||||||||||:|| ||||  |::|||||||  |||||||   ||||| :|:
Db        243 VVGNITKAVSSYILAQRVAAGLSSTGPVAALIASTVSLAISPLAFAGIADKFNHAKSLES 302

Qy        312 FAKQFRKFGYDGDHLLAEYQRGVGTIEASLTTISTALGAVSAGVSAAAVGSAVGTPIALL 371
              :|::|:| ||||:||||||| |||:||:| |:||| |:: |||||| || : :||||
Db        303 YAERFKKLGYDGDNLLAEYQRGTGTIDASVTAINTALAAIAGGVSAAAAGSVIASPIALL 362

Qy        372 VAGVTGLISGILEASKQAMFESVANRLQGKILEWEKQNGGQNYFDKGYDSRYAAYLANNL 431
              |:|:||:|| ||: |||||| |||::  ||:|||| | |:|||: |||:||  | | :|:
Db        363 VSGITGVISTILQYSKQAMFEHVANKIHNKIVEWEKNNHGKNYFENGYDARYLANLQDNM 422

Qy        432 KFLSELNKELEAERVIAITQQRWDNNIGELAGITKLGERIKSGKAYADAFEDGKKVEAGS 491
              |||  |||||:|||||||||:|||||:||||::|||:: ||||| ||||:|| ::|
Db        423 KFLLNLNKELQAERVIAITQQQWDNNIGDLAGISRLGEKVLSGKAYVDAFEEGKHIKADK 482

Qy        492 NITLDAKTGIIDISNSNGKKTQALHFTSPLLTAGTESRERLTNGKYSYINKLKFGRVKNW 551
               : ||:  ||||:|||    |||  : | :|||| ||| |||:   ||| ||  ||  || :|
Db        483 LVQLDSANGIIDVSNSGKAKTQHILFRTPLLTPGTEHRERVQTGKYEYITKLNINRVDSW 542

Qy        552 QVTDGEASSKLDFSKVIQRV-------AETEGTDEIGLIVNAKAGNDDIFVGQGKMNIDG 604
              ::||| |||   | : |:||:          | |   :|     |:|::||| |   |||
Db        543 KITDGAASSTFDLTNVVQRIGIELDNAGNVTKTKETKIIAKLGEGDDNVFVGSGTTEIDG 602

Qy        605 GDGHDRVFYSKDGGFGNITVDGTSATEAGSYTVNRKVARGDIYHEVVKRQETKVGKRTET 664
              |:|:||| ||: | :| :|:| |  || |||||||  |   |||       || | |
Db        603 GEGYDRVHYSR-GNYGALTIDATKETEQGSYTVNRFVETGKALHEVTSTHTALVGNREEK 661

Qy        665 IQYRDYELRKVGYGYQSTDNLKSVEEVIGSQFNDVFKGSKFNDIFHSGEGDDLLDGGAGD 724
              |:|| :    :   || : | ||:|||:||:  ||:||||||||| |: |:| | :|| |:
Db        662 IEYR-HSNNQHHAGYYTKDTLKAVEEIIGTSHNDIFKGSKFNDAFNGGDGVDTIDGNDGN 720

Qy        725 DRLFGGKGNDRLSGDEGDDLLDGGSGDDVLNGGAGNDVYIFRKGDGNDTLYDGTGNDKLA 784
              ||||||||:| |  |  ||| :|||  |:|:|:|| |:|:::  ||||||  :  |  |||||:
Db        721 DRLFGGKGDDILDGGNGDDFIDGGKGNDLLHGGKGDDIFVHRKGDGNDIITDSDGNDKLS 780

Qy        785 FADANISDIMIERTKEGIIVKRNDHSGSINIPRWY----ITSNLQNYQSNKTDHKIEQLI 840
              |:|:|: |: |: |:  |  :::  |       :  |  |:       : ||:: |  | |||::|
Db        781 FSDSNLKDLTFEKVKHNLVI-TNSKKEKVTIQNWFREADFAKEVPNYKATK-DEKIEEII 838

Qy        841 GKDGSYITSDQIDKILQDKKDGTVITSQELKKLADENKSQKLSASDIASSLNKLVGSMAL 900
              |::|  ||| |:|  ::  |   |     ||  ||  |:  |  :   |  |  :: :||:||: |::
Db        839 GQNGERITSKQVDDLI--AKGNGKITQDELSKVVDNYELLKHS-KNVTNSLDKLISSVSA 895

Qy        901 FGTANSVSSNALQPITQPTQGI 922
              | ::|   :  : | :   |  :
Db        896 FTSSNDSRNVLVAPTSMLDQSL 917


RESULT 8
AAR42380
ID   AAR42380 standard; Protein; 924 AA.
XX
AC   AAR42380;
XX
DT   25-MAR-2003  (updated)
DT   19-APR-1994  (first entry)
XX
DE   Recombinant leukotoxin peptide (split) from plasmid pGCH4.
XX
KW   Haemophilus somnus; immunogenic; haemolysin; LppB; LppC;
KW   thromboembolic meningoencephalitis; septicaemia; arthritis;
KW   pneumonia; lktA gene; haemin-binding protein; fusion protein.
XX
OS   Pasteurella haemolytica.
XX
PN   WO9321323-A1.
XX
PD   28-OCT-1993.
XX
PF   05-APR-1993;   93WO-CA00135.
XX
PR   09-APR-1992;   92US-0865050.
PR   04-JUN-1992;   92US-0893424.
PR   04-JUN-1992;   92US-0893426.
PR   29-MAR-1993;   93US-0038287.
PR   29-MAR-1993;   93US-0038288.
PR   29-MAR-1993;   93US-0038719.
XX
PA   (UYSA-) UNIV SASKATCHEWAN.
XX
PI   Harland RJ,  Pfeiffer CG,  Pontarollo RA,  Potter AA;
PI    Rioux C,  Theisen M;
XX
DR   WPI; 1993-351733/44.
DR   N-PSDB; AAQ51082.
XX
PT   Haemophilus somnus immunogenic proteins used in vaccines -
PT   selected from haemin-binding protein, haemolysin, LppB and LppC,
PT   and corresp. DNA
XX
PS   Disclosure; Fig 6; 119pp; English.
XX
CC   The hmb gene encoding the haemin-binding protein was expressed in
CC   E. coli as a fusion to the Pasteurella haemolytica leukotoxin gene
CC   lktA coded for by plasmid pAA352.  The hmb gene fragment was taken
CC   from pRAP504 and starts at the codon for the 33 rd amino acid residue
CC   of ORF1.  The haemin binding protein can be used in vaccines for
CC   preventing or treating H. somnus infections, which cause thromboembolic
CC   meningo-encephalitis, septicaemia, arthritis and pneumonia in
CC   vertebrates.
CC   See also AAR42370-86.
CC   (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ   Sequence   924 AA;
```

```
Query Match             50.1%;  Score 2326.5;  DB 14;  Length 924;
Best Local Similarity   50.0%;  Pred. No. 1.7e-136;
Matches  461;  Conservative  174;  Mismatches  262;  Indels   25;  Gaps   12;

Qy         15 LNSTKSGLKNLYLAIPKD--YDPQKGGTLNDFIKAADELGIARLAEEPNHTETAKKSVDT 72
Db          7 LSFPKTGAKKIILYIPQNYQYDTEQGNGLQDLVKAAEELGIEVQREERNNIATAQTSLGT 66

Qy         73 VNQFLSLTQTGIAISATKLEKFLQKHSTNKLAKGLDSVENIDRKLGKASNVLSTLSSFLG 132
Db         67 IQTAIGLTERGIVLSAPQIDKLLQK---TKAGQALGSAESIVQNANKAKTVLSGIQSILG 123

Qy        133 TALAGIELDSLIKKGDAAPDALAKASIDLINEIIGNLSQSTQTIEAFSSQLAKLGSTISQ 192
Db        124 SVLAGMDLDEAL-QNNSNQHALAKAGLELTNSLIENIANSVKTLDEFGEQISQFGSKLQN 182

Qy        193 AKGFSNIGNKLQNL-NFSKTNLGLEIITGLLSGISAGFALADKNASTGKKVAAGFELSNQ 251
Db        183 IKGLGTLGDKLKNIGGLDKAGLGLDVISGLLSGATAALVLADKNASTAKKVGAGFELANQ 242

Qy        252 VIGNVTKAISSYVLAQRVAAGLSTTGAVAALITSSIMLAISPLAFMNAADKFNHANALDE 311
Db        243 VVGNITKAVSSYILAQRVAAGLSSTGPVAALIASTVSLAISPLAFAGIADKFNHAKSLES 302

Qy        312 FAKQFRKFGYDGDHLLAEYQRGVGTIEASLTTISTALGAVSAGVSAAAVGSAVGTPIALL 371
Db        303 YAERFKKLGYDGDNLLAEYQRGTGTIDASVTAINTALAAIAGGVSAAAAGSVIASPIALL 362

Qy        372 VAGVTGLISGILEASKQAMFESVANRLQGKILEWEKQNGGQNYFDKGYDSRYAAYLANNL 431
Db        363 VSGITGVISTILQYSKQAMFEHVANKIHNKIVEWEKNNHGKNYFENGYDARYLANLQDNM 422

Qy        432 KFLSELNKELEAERVIAITQQRWDNNIGELAGITKLGERIKSGKAYADAFEDGKKVEAGS 491
Db        423 KFLLNLNKELQAERVIAITQQQWDNNIGDLAGISRLGEKVLSGKAYVDAFEEGKHIKADK 482

Qy        492 NITLDAKTGIIDISNSNGKKTQALHFTSPLLTAGTESRERLTNGKYSYINKLKFGRVKNW 551
Db        483 LVQLDSANGIIDVSNSGKAKTQHILFRTPLLTPGTEHRERVQTGKYEYITKLNINRVDSW 542

Qy        552 QVTDGEASSKLDFSKVIQRV-------AETEGTDEIGLIVNAKAGNDDIFVGQGKMNIDG 604
Db        543 KITDGAASSTFDLTNVVQRIGIELDNAGNVTKTKETKIIAKLGEGDDNVFVGSGTTEIDG 602

Qy        605 GDGHDRVFYSKDGGFGNITVDGTSATEAGSYTVNRKVARGDIYHEVVKRQETKVGKRTET 664
Db        603 GEGYDRVHYSR-GNYGALTIDATKETEQGSYTVNRFVETGKALHEVTSTHTALVGNREEK 661

Qy        665 IQYRDYELRKVGYGYQSTDNLKSVEEVIGSQFNDVFKGSKFNDIFHSGEGDDLLDGGAGD 724
Db        662 IEYR-HSNNQHHAGYYTKDTLKAVEEIIGTSHNDIFKGSKFNDAFNGGDGVDTIDGNDGN 720

Qy        725 DRLFGGKGNDRLSGDEGDDLLDGGSGDDVLNGGAGNDVYIFRKGDGNDTLYDGTGNDKLA 784
Db        721 DRLFGGKGDDILDGGNGDDFIDGGKGNDLLHGGKGDDIFVHRKGDGNDIITDSDGNDKLS 780

Qy        785 FADANISDIMIERTKEGIIVKRNDHSGSINIPRWY----ITSNLQNYQSNKTDHKIEQLI 840
Db        781 FSDSNLKDLTFEKVKHNLVI-TNSKKEKVTIQNWFREADFAKEVPNYKATK-DEKIEEII 838

Qy        841 GKDGSYITSDQIDKILQDKKDGTVITSQELKKLADENKSQKLSASDIASSLNKLVGSMAL 900
Db        839 GQNGERITSKQVDDLI--AKGNGKITQDELSKVVDNYELLKHS-KNVTNSLDKLISSVSA 895

Qy        901 FGTANSVSSNALQPITQPTQGI 922
Db        896 FTSSNDSRNVLVAPTSMLDQSL 917
```

RESULT 9
AAR42378
ID   AAR42378 standard; Protein; 924 AA.
XX
AC   AAR42378;
XX
DT   25-MAR-2003 (updated)
DT   19-APR-1994 (first entry)
XX
DE   Recombinant leukotoxin peptide (split) from plasmid pGCH5.
XX
KW   Haemophilus somnus; immunogenic; haemolysin; LppB; LppC;
KW   thromboembolic meningoencephalitis; septicaemia; arthritis;
KW   pneumonia; lktA gene; haemin-binding protein; fusion protein.
XX
OS   Pasteurella haemolytica.
XX
PN   WO9321323-A1.
XX
PD   28-OCT-1993.
XX
PF   05-APR-1993;   93WO-CA00135.
XX
PR   09-APR-1992;   92US-0865050.
PR   04-JUN-1992;   92US-0893424.
PR   04-JUN-1992;   92US-0893426.
PR   29-MAR-1993;   93US-0038287.
PR   29-MAR-1993;   93US-0038288.
PR   29-MAR-1993;   93US-0038719.
XX
PA   (UYSA-) UNIV SASKATCHEWAN.
XX
PI   Harland RJ, Pfeiffer CG, Pontarollo RA, Potter AA;
PI   Rioux C, Theisen M;
XX
DR   WPI; 1993-351733/44.
DR   N-PSDB; AAQ51081.
XX
PT   Haemophilus somnus immunogenic proteins used in vaccines -
PT   selected from haemin-binding protein, haemolysin, LppB and LppC,
PT   and corresp. DNA
XX
PS   Disclosure; Fig 5; 119pp; English.
XX
CC   The hmb gene encoding the haemin-binding protein was expressed in
CC   E. coli as a fusion to the Pasteurella haemolytica leukotoxin gene
CC   lktA coded for by plasmid pAA352. The hmb gene fragment was taken
CC   from pRAP501 and starts at the codon for the third amino acid residue
CC   of ORF1. The haemin binding protein can be used in vaccines for
CC   preventing or treating H. somnus infections, which cause thromboembolic
CC   meningo-encephalitis, septicaemia, arthritis and pneumonia in
CC   vertebrates.
CC   See also AAR42370-86.
CC   (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ   Sequence   924 AA;

```
Query Match             50.1%;  Score 2326.5;  DB 14;  Length 924;
Best Local Similarity   50.0%;  Pred. No. 1.7e-136;
Matches  461;  Conservative  174;  Mismatches  262;  Indels   25;  Gaps   12;

Qy         15 LNSTKSGLKNLYLAIPKD--YDPQKGGTLNDFIKAADELGIARLAEEPNHTETAKKSVDT 72
Db          7 LSFPKTGAKKIILYIPQNYQYDTEQGNGLQDLVKAAEELGIEVQREERNNIATAQTSLGT 66

Qy         73 VNQFLSLTQTGIAISATKLEKFLQKHSTNKLAKGLDSVENIDRKLGKASNVLSTLSSFLG 132
Db         67 IQTAIGLTERGIVLSAPQIDKLLQK---TKAGQALGSAESIVQNANKAKTVLSGIQSILG 123

Qy        133 TALAGIELDSLIKKGDAAPDALAKASIDLINEIIGNLSQSTQTIEAFSSQLAKLGSTISQ 192
Db        124 SVLAGMDLDEAL-QNNSNQHALAKAGLELTNSLIENIANSVKTLDEFGEQISQFGSKLQN 182

Qy        193 AKGFSNIGNKLQNL-NFSKTNLGLEIITGLLSGISAGFALADKNASTGKKVAAGFELSNQ 251
Db        183 IKGLGTLGDKLKNIGGLDKAGLGLDVISGLLSGATAALVLADKNASTAKKVGAGFELANQ 242
```

```
Qy        252 VIGNVTKAISSYVLAQRVAAGLSTTGAVAALITSSIMLAISPLAFMNAADKFNHANALDE 311
Db        243 VVGNITKAVSSYILAQRVAAGLSSTGPVAALIASTVSLAISPLAFAGIADKFNHAKSLES 302

Qy        312 FAKQFRKFGYDGDHLLAEYQRGVGTIEASLTTISTALGAVSAGVSAAAVGSAVGTPIALL 371
Db        303 YAERFKKLGYDGDNLLAEYQRGTGTIDASVTAINTALAAIAGGVSAAAAGSVIASPIALL 362

Qy        372 VAGVTGLISGILEASKQAMFESVANRLQGKILEWEKQNGGQNYFDKGYDSRYAAYLANNL 431
Db        363 VSGITGVISTILQYSKQAMFEHVANKIHNKIVEWEKNNHGKNYFENGYDARYLANLQDNM 422

Qy        432 KFLSELNKELEAERVIAITQQRWDNNIGELAGITKLGERIKSGKAYADAFEDGKKVEAGS 491
Db        423 KFLLNLNKELQAERVIAITQQQWDNNIGDLAGISRLGEKVLSGKAYVDAFEEGKHIKADK 482

Qy        492 NITLDAKTGIIDISNSNGKKTQALHFTSPLLTAGTESRERLTNGKYSYINKLKFGRVKNW 551
Db        483 LVQLDSANGIIDVSNSGKAKTQHILFRTPLLTPGTEHRERVQTGKYEYITKLNINRVDSW 542

Qy        552 QVTDGEASSKLDFSKVIQRV-------AETEGTDEIGLIVNAKAGNDDIFVGQGKMNIDG 604
Db        543 KITDGAASSTFDLTNVVQRIGIELDNAGNVTKTKETKIIAKLGEGDDNVFVGSGTTEIDG 602

Qy        605 GDGHDRVFYSKDGGFGNITVDGTSATEAGSYTVNRKVARGDIYHEVVKRQETKVGKRTET 664
Db        603 GEGYDRVHYSR-GNYGALTIDATKETEQGSYTVNRFVETGKALHEVTSTHTALVGNREEK 661

Qy        665 IQYRDYELRKVGYGYQSTDNLKSVEEVIGSQFNDVFKGSKFNDIFHSGEGDDLLDGGAGD 724
Db        662 IEYR-HSNNQHHAGYYTKDTLKAVEEIIGTSHNDIFKGSKFNDAFNGGDGVDTIDGNDGN 720

Qy        725 DRLFGGKGNDRLSGDEGDDLLDGGSGDDVLNGGAGNDVYIFRKGDGNDTLYDGTGNDKLA 784
Db        721 DRLFGGKGDDILDGGNGDDFIDGGKGNDLLHGGKGDDIFVHRKGDGNDIITDSDGNDKLS 780

Qy        785 FADANISDIMIERTKEGIIVKRNDHSGSINIPRWY----ITSNLQNYQSNKTDHKIEQLI 840
Db        781 FSDSNLKDLTFEKVKHNLVI-TNSKKEKVTIQNWFREADFAKEVPNYKATK-DEKIEEII 838

Qy        841 GKDGSYITSDQIDKILQDKKDGTVITSQELKKLADENKSQKLSASDIASSLNKLVGSMAL 900
Db        839 GQNGERITSKQVDDLI--AKGNGKITQDELSKVVDNYELLKHS-KNVTNSLDKLISSVSA 895

Qy        901 FGTANSVSSNALQPITQPTQGI 922
Db        896 FTSSNDSRNVLVAPTSMLDQSL 917
```

RESULT 10
AAR14482
ID   AAR14482 standard; Protein; 926 AA.
XX
AC   AAR14482;
XX
DT   15-JAN-1992 (first entry)
XX
DE   LKT352.
XX
KW   Antigen; leukotoxin; vaccine; lktA.
XX
OS   Pasteurella haemolytica.
XX
PN   WO9115237-A.
XX
PD   17-OCT-1991.
XX
PF   17-OCT-1991;   91WO-CA00170.
XX
PR   05-APR-1990;   90US-0504850.
XX
PA   (UYSA-) UNIV SASKATCHEWAN.
XX
PI   Acres SD, Bariuk LA, Potter AA, Lawman MJP;
XX
DR   WPI; 1991-324967/44.
XX
PT   Vaccines for Pasteurella haemolytica infection in cattle -
PT   comprise sub-unit antigens from P haemolytica fimbrial protein,
PT   plasmin receptor, 50 K outer membrane protein and leukotoxin.
XX
PS   Disclosure; Fig 5; 92pp; English.
XX
CC   LKT352 is 98% homologous with authentic leukotoxin and migrates
CC   to the same position on gels.
CC   The LKT352 gene was prepd. as follows: lktA, an MaeI fragment
CC   contg. the gene was ligated into the SmaI site of pUC13 to form
CC   pAA179. From this, two constructs were made in the ptac-based
CC   vector, pGH432:lacI digested with SmaI. One, pAA342, consisted of
CC   the 5' AhaIII fragment from lktA while the other, pAA345, contained
CC   the entire MaeI fragment. Clone pAA342 expressed a truncated
CC   leukotoxin peptide at high levels while pAA345 expressed full
CC   length leukotoxin at very low levels. The 3' end of the lktA gene
CC   of pAA345 was therefore ligated to StyI/BamHI digested pAA342 to
CC   yield pAA352 contg. the LKT352 sequence. The protein expressed
CC   from the vector can be used to prepare a subunit vaccine with
CC   other P. haemolytica antigens, e.g. fimbrial protein, plasmin
CC   receptor or 50K outer membrane protein. The vaccines can be used
CC   to protect cattle from respiratory diseases such as pneumonia, esp.
CC   shipping fever pneumonia.
CC   See also AAR14481, 83,84 and 85.
XX
SQ   Sequence   926 AA;

Query Match 50.1%; Score 2326.5; DB 12; Length 926;
Best Local Similarity 50.0%; Pred. No. 1.7e-136;
Matches 461; Conservative 174; Mismatches 262; Indels 25; Gaps 12;

```
Qy         15 LNSTKSGLKNLYLAIPKD--YDPQKGGTLNDFIKAADELGIARLAEEPNHTETAKKSVDT 72
Db          7 LSFPKTGAKKIILYIPQNYQYDTEQGNGLQDLVKAAEELGIEVQREERNNIATAQTSLGT 66

Qy         73 VNQFLSLTQTGIAISATKLEKFLQKHSTNKLAKGLDSVENIDRKLGKASNVLSTLSSFLG 132
Db         67 IQTAIGLTERGIVLSAPQIDKLLQK---TKAGQALGSAESIVQNANKAKTVLSGIQSILG 123

Qy        133 TALAGIELDSLIKKGDAAPDALAKASIDLINEIIGNLSQSTQTIEAFSSQLAKLGSTISQ 192
Db        124 SVLAGMDLDEAL-QNNSNQHALAKAGLELTNSLIENIANSVKTLDEFGEQISQFGSKLQN 182

Qy        193 AKGFSNIGNKLQNL-NFSKTNLGLEIITGLLSGISAGFALADKNASTGKKVAAGFELSNQ 251
Db        183 IKGLGTLGDKLKNIGGLDKAGLGLDVISGLLSGATAALVLADKNASTAKKVGAGFELANQ 242

Qy        252 VIGNVTKAISSYVLAQRVAAGLSTTGAVAALITSSIMLAISPLAFMNAADKFNHANALDE 311
Db        243 VVGNITKAVSSYILAQRVAAGLSSTGPVAALIASTVSLAISPLAFAGIADKFNHAKSLES 302

Qy        312 FAKQFRKFGYDGDHLLAEYQRGVGTIEASLTTISTALGAVSAGVSAAAVGSAVGTPIALL 371
Db        303 YAERFKKLGYDGDNLLAEYQRGTGTIDASVTAINTALAAIAGGVSAAAAGSVIASPIALL 362

Qy        372 VAGVTGLISGILEASKQAMFESVANRLQGKILEWEKQNGGQNYFDKGYDSRYAAYLANNL 431
Db        363 VSGITGVISTILQYSKQAMFEHVANKIHNKIVEWEKNNHGKNYFENGYDARYLANLQDNM 422

Qy        432 KFLSELNKELEAERVIAITQQRWDNNIGELAGITKLGERIKSGKAYADAFEDGKKVEAGS 491
Db        423 KFLLNLNKELQAERVIAITQQQWDNNIGDLAGISRLGEKVLSGKAYVDAFEEGKHIKADK 482

Qy        492 NITLDAKTGIIDISNSNGKKTQALHFTSPLLTAGTESRERLTNGKYSYINKLKFGRVKNW 551
Db        483 LVQLDSANGIIDVSNSGKAKTQHILFRTPLLTPGTEHRERVQTGKYEYITKLNINRVDSW 542

Qy        552 QVTDGEASSKLDFSKVIQRV-------AETEGTDEIGLIVNAKAGNDDIFVGQGKMNIDG 604
```

```
Db        543 KITDGAASSTFDLTNVVQRIGIELDNAGNVTKTKETKIIAKLGEGDDNVFVGSGTTEIDG 602

Qy        605 GDGHDRVFYSKDGGFGNITVDGTSATEAGSYTVNRKVARGDIYHEVVKRQETKVGKRTET 664
              |:|:||| ||: | :| :|:| |  || ||||||| |   |   |||         || | |
Db        603 GEGYDRVHYSR-GNYGALTIDATKETEQGSYTVNRFVETGKALHEVTSTHTALVGNREEK 661

Qy        665 IQYRDYELRKVGYGYQSTDNLKSVEEVIGSQFNDVFKGSKFNDIFHSGEGDDLLDGGAGD 724
              |:|| :    :     || : | ||:|||:||:  ||:||||||||| |: |:| | :|| |:
Db        662 IEYR-HSNNQHHAGYYTKDTLKAVEEIIGTSHNDIFKGSKFNDAFNGGDGVDTIDGNDGN 720

Qy        725 DRLFGGKGNDRLSGDEGDDLLDGGSGDDVLNGGAGNDVYIFRKGDGNDTLYDGTGNDKLA 784
              ||||||||:| | |  ||| :||| |:|:|:|| |:|::: |||||| : |  |||||:
Db        721 DRLFGGKGDDILDGGNGDDFIDGGKGNDLLHGGKGDDIFVHRKGDGNDIITDSDGNDKLS 780

Qy        785 FADANISDIMIERTKEGIIVKRNDHSGSINIPRWY----ITSNLQNYQSNKTDHKIEQLI 840
              |:|:|: |: |: |  ::: |        :  | |:        : ||:: | | |||::|
Db        781 FSDSNLKDLTFEKVKHNLVI-TNSKKEKVTIQNWFREADFAKEVPNYKATK-DEKIEEII 838

Qy        841 GKDGSYITSDQIDKILQDKKDGTVITSQELKKLADENKSQKLSASDIASSLNKLVGSMAL 900
              |::| ||| |:| ::    |        ||  || |: |  :  | |  :: :||:||: |::
Db        839 GQNGERITSKQVDDLI--AKGNGKITQDELSKVVDNYELLKHS-KNVTNSLDKLISSVSA 895

Qy        901 FGTANSVSSNALQPITQPTQGI 922
              | ::|   :  : | :   | :
Db        896 FTSSNDSRNVLVAPTSMLDQSL 917
```

RESULT 11

```
AAR34545
ID   AAR34545 standard; Protein; 926 AA.
XX
AC   AAR34545;
XX
DT   25-MAR-2003  (updated)
DT   23-AUG-1993  (first entry)
XX
DE   Leukotoxin 352 produced from pAA352.
XX
KW   Vector; LKT 352; flanking; recombinant; antigen; somatostatin;
KW   gonadotropin releasing hormone; rotavirus viral protein 4;
KW   carrier protein; lactation; reproduction.
XX
OS   Pasteurella haemolytica.
XX
PN   WO9308290-A1.
XX
PD   29-APR-1993.
XX
PF   15-OCT-1992;   92WO-CA00449.
XX
PR   16-OCT-1991;   91US-0779171.
PR   14-OCT-1992;   92US-0960932.
XX
PA   (UYSA-) UNIV SASKATCHEWAN.
XX
PI   Hughes HPA,  Potter AA,  Redmond MJ;
XX
DR   WPI; 1993-152482/18.
DR   N-PSDB; AAQ41317.
XX
PT   Immunological carrier system with enhanced immunogenicity -
PT   comprises chimeric protein comprising leuco:toxin peptide or
PT   homologous protein fused to antigen esp. somatostatin or
PT   gonadotropin releasing hormone
XX
PS   Disclosure; Fig 3; 95pp; English.
XX
CC   Gene libraries of P. haemolytica A1 (strain B122) were constructed
CC   in lambda gt11 and pUC13.  Resulting clones were used to transform E.
CC   coli and individual colonies were pooled and screened for reaction
CC   with serum from a calf which had survived a P. haemolytica infection
CC   and that had been boosted with a conc. culture supernatant of P.
CC   haemolytica to increase anti-leukotoxin antibody levels.  Positive
CC   colonies were screened for their ability to produce leukotoxin by
CC   incubating cell lysates with bovine neutrophils and measuring the
CC   release of lactate dehydrogenase from the neutrophils.  A 4kb
CC   fragment was obtd.  Progressively larger clones were isolated by
CC   chromosome walking to isolate full length recombinants of ca. 8kb,
CC   in pAA114.  The clone was subjected to restriction enzyme digestion
CC   to yield two clones, one expressing truncated leukotoxin peptide at
CC   high levels and the other expressing the full length leukotoxin at
CC   low levels.  The 3' end of the lktA gene from the full length clone
CC   was ligated to the truncated gene clone to yield plasmid pAA352.  The
CC   clone was used to produce chimeric proteins by gene fusion with an
CC   antigen coding sequence, e.g. the coding sequence of somatostatin,
CC   gonadotrophin releasing hormone or rotavirus viral protein 4, i.e.
CC   leukotoxin works as a carrier protein to bring about a larger
CC   immune response than the antigen alone.  Immunisation with these
CC   antigens can regulate growth rate, lactation and reproductive
CC   efficiency. See also AAR34546-8.
CC   (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ   Sequence   926 AA;

  Query Match             50.1%;  Score 2326.5;  DB 14;  Length 926;
  Best Local Similarity   50.0%;  Pred. No. 1.7e-136;
  Matches  461;  Conservative  174;  Mismatches  262;  Indels   25;  Gaps   12;

Qy         15 LNSTKSGLKNLYLAIPKD--YDPQKGGTLNDFIKAADELGIARLAEEPNHTETAKKSVDT 72
              |:  |:| | : | ||::  || ::|  | | :|||:||||      || |:  ||: |: |
Db          7 LSFPKTGAKKIILYIPQNYQYDTEQGNGLQDLVKAAEELGIEVQREERNNIATAQTSLGT 66

Qy         73 VNQFLSLTQTGIAISATKLEKFLQKHSTNKLAKGLDSVENIDRKLGKASNVLSTLSSFLG 132
              :   : ||: || :|| :::| |||     |  : | | |:| :     ||  ||| : | ||
Db         67 IQTAIGLTERGIVLSAPQIDKLLQK---TKAGQALGSAESIVQNANKAKTVLSGIQSILG 123

Qy        133 TALAGIELDSLIKKGDAAPDALAKASIDLINEIIGNLSQSTQTIEAFSSQLAKLGSTISQ 192
              : |||::||  : : ::   ||||  ::| | :| |:: | :|:: |   |::: || :
Db        124 SVLAGMDLDEAL-QNNSNQHALAKAGLELTNSLIENIANSVKTLDEFGEQISQFGSKLQN 182

Qy        193 AKGFSNIGNKLQNL-NFSKTNLGLEIITGLLSGISAGFALADKNASTGKKVAAGFELSNQ 251
              ||   :|:||:|:    |  |||::|:||||| :|    ||||||| ||| |||||:||
Db        183 IKGLGTLGDKLKNIGGLDKAGLGLDVISGLLSGATAALVLADKNASTAKKVGAGFELANQ 242

Qy        252 VIGNVTKAISSYVLAQRVAAGLSTTGAVAALITSSIMLAISPLAFMNAADKFNHANALDE 311
              |:||:|||:|||:||||||||||:|| |||||  |:: ||||||||    |||||| :|:
Db        243 VVGNITKAVSSYILAQRVAAGLSSTGPVAALIASTVSLAISPLAFAGIADKFNHAKSLES 302

Qy        312 FAKQFRKFGYDGDHLLAEYQRGVGTIEASLTTISTALGAVSAGVSAAAVGSAVGTPIALL 371
              :|::|:| |||||:||||||||| |||:||:| |:||| |:: ||||| || : :|||||
Db        303 YAERFKKLGYDGDNLLAEYQRGTGTIDASVTAINTALAAIAGGVSAAAAGSVIASPIALL 362

Qy        372 VAGVTGLISGILEASKQAMFESVANRLQGKILEWEKQNGGQNYFDKGYDSRYAAYLANNL 431
              |:|:||:|| ||: |||||| ||||::  ||:||| |:|||  ||:|||: |||:|| | :|:
Db        363 VSGITGVISTILQYSKQAMFEHVANKIHNKIVEWEKNNHGKNYFENGYDARYLANLQDNM 422

Qy        432 KFLSELNKELEAERVIAITQQRWDNNIGELAGITKLGERIKSGKAYADAFEDGKKVEAGS 491
              |||  |||||:||||||||||:||||||:|||::||::| |||| ||||:|| ::|
Db        423 KFLLNLNKELQAERVIAITQQQWDNNIGDLAGISRLGEKVLSGKAYVDAFEEGKHIKADK 482

Qy        492 NITLDAKTGIIDISNSNGKKTQALHFTSPLLTAGTESRERLTNGKYSYINKLKFGRVKNW 551
               : ||:  ||||:|||   ||| : | :|||| ||| |||:  ||| || ||   || :|
Db        483 LVQLDSANGIIDVSNSGKAKTQHILFRTPLLTPGTEHRERVQTGKYEYITKLNINRVDSW 542

Qy        552 QVTDGEASSKLDFSKVIQRV-------AETEGTDEIGLIVNAKAGNDDIFVGQGKMNIDG 604
              ::||| ||| | : |:||:            | | :|     |:|::||| |    |||
Db        543 KITDGAASSTFDLTNVVQRIGIELDNAGNVTKTKETKIIAKLGEGDDNVFVGSGTTEIDG 602

Qy        605 GDGHDRVFYSKDGGFGNITVDGTSATEAGSYTVNRKVARGDIYHEVVKRQETKVGKRTET 664
              |:|:||| ||: | :| :|:| |  || ||||||| |   |   |||         || | |
Db        603 GEGYDRVHYSR-GNYGALTIDATKETEQGSYTVNRFVETGKALHEVTSTHTALVGNREEK 661

Qy        665 IQYRDYELRKVGYGYQSTDNLKSVEEVIGSQFNDVFKGSKFNDIFHSGEGDDLLDGGAGD 724
              |:|| :    :     || : | ||:|||:||:  ||:||||||||| |: |:| | :|| |:
Db        662 IEYR-HSNNQHHAGYYTKDTLKAVEEIIGTSHNDIFKGSKFNDAFNGGDGVDTIDGNDGN 720
```

```
Qy        725 DRLFGGKGNDRLSGDEGDDLLDGGSGDDVLNGGAGNDVYIFRKGDGNDTLYDGTGNDKLA 784
Db        721 DRLFGGKGDDILDGGNGDDFIDGGKGNDLLHGGKGDDIFVHRKGDGNDIITDSDGNDKLS 780

Qy        785 FADANISDIMIERTKEGIIVKRNDHSGSINIPRWY----ITSNLQNYQSNKTDHKIEQLI 840
Db        781 FSDSNLKDLTFEKVKHNLVI-TNSKKEKVTIQNWFREADFAKEVPNYKATK-DEKIEEII 838

Qy        841 GKDGSYITSDQIDKILQDKKDGTVITSQELKKLADENKSQKLSASDIASSLNKLVGSMAL 900
Db        839 GQNGERITSKQVDDLI--AKGNGKITQDELSKVVDNYELLKHS-KNVTNSLDKLISSVSA 895

Qy        901 FGTANSVSSNALQPITQPTQGI 922
Db        896 FTSSNDSRNVLVAPTSMLDQSL 917
```

```
RESULT 12
AAR50291
ID   AAR50291 standard; Protein; 926 AA.
XX
AC   AAR50291;
XX
DT   06-OCT-1994  (first entry)
XX
DE   Recombinant leukotoxin from plasmid pAA352.
XX
KW   Vaccine; outer membrane protein; OMP; Haemophilus somnus;
KW   iron regulated protein; leukotoxin; Pasteurella haemolytica;
KW   LKT352.
XX
OS   Pasteurella haemolytica A1 (strain B122).
XX
PN   CA2099707-A.
XX
PD   03-JAN-1994.
XX
PF   29-JUN-1993;   93CA-2099707.
XX
PR   02-JUL-1992;   92US-0908253.
XX
PA   (UYSA-) UNIV SASKATCHEWAN.
XX
PI   Harland RJ,  Potter AA;
XX
DR   WPI; 1994-092909/12.
DR   N-PSDB; AAQ44760.
XX
PT   Haemophilus somnus outer membrane protein extract -
PT   enriched with iron-regulated proteins, opt. contg.
PT   leuco:toxin antigens, for use as vaccine
XX
PS   Claim 5; Fig 5; 78pp; English.
XX
CC   A vaccine comprising an outer membrane protein (OMP) extract of
CC   Haemophilus somnus enriched with iron regulated proteins is new.
CC   The vaccine pref. further comprises an immunogenic leukotoxin
CC   polypeptide, esp. an immunogenic Pasteurella haemolytica leukotoxin
CC   homologous to LKT352.  Example 1.2 describes the prodn. of
CC   P. haemolytica recombinant leukotoxin from pAA352.
CC   Two expression constructs were made.  One, pAA342, contained the
CC   5'-AhaII fragment of the ltkA gene, while the other, pAA345,
CC   contained the entire lktA gene.  pAA342 expressed a truncated
CC   leukotoxin peptide at high levels, while pAA345 expressed full
CC   length leukotoxin at very low levels.  Therefore, the 3' end
CC   of the lktA gene was ligated into pAA342, yielding plasmid pAA352.
CC   LKT352 or new leukotoxin is 98% homologous to authentic
CC   leukotoxin.
CC   NB: the protein sequence in Fig 5 comprises 926 amino acids,
CC   however this protein is described in the text as having
CC   931 amino acids.
XX
SQ   Sequence   926 AA;
```

```
  Query Match             50.1%;  Score 2326.5;  DB 15;  Length 926;
  Best Local Similarity   50.0%;  Pred. No. 1.7e-136;
  Matches  461;  Conservative  174;  Mismatches  262;  Indels   25;  Gaps   12;

Qy         15 LNSTKSGLKNLYLAIPKD--YDPQKGGTLNDFIKAADELGIARLAEEPNHTETAKKSVDT 72
Db          7 LSFPKTGAKKIILYIPQNYQYDTEQGNGLQDLVKAAEELGIEVQREERNNIATAQTSLGT 66

Qy         73 VNQFLSLTQTGIAISATKLEKFLQKHSTNKLAKGLDSVENIDRKLGKASNVLSTLSSFLG 132
Db         67 IQTAIGLTERGIVLSAPQIDKLLQK---TKAGQALGSAESIVQNANKAKTVLSGIQSILG 123

Qy        133 TALAGIELDSLIKKGDAAPDALAKASIDLINEIIGNLSQSTQTIEAFSSQLAKLGSTISQ 192
Db        124 SVLAGMDLDEAL-QNNSNQHALAKAGLELTNSLIENIANSVKTLDEFGEQISQFGSKLQN 182

Qy        193 AKGFSNIGNKLQNL-NFSKTNLGLEIITGLLSGISAGFALADKNASTGKKVAAGFELSNQ 251
Db        183 IKGLGTLGDKLKNIGGLDKAGLGLDVISGLLSGATAALVLADKNASTAKKVGAGFELANQ 242

Qy        252 VIGNVTKAISSYVLAQRVAAGLSTTGAVAALITSSIMLAISPLAFMNAADKFNHANALDE 311
Db        243 VVGNITKAVSSYILAQRVAAGLSSTGPVAALIASTVSLAISPLAFAGIADKFNHAKSLES 302

Qy        312 FAKQFRKFGYDGDHLLAEYQRGVGTIEASLTTISTALGAVSAGVSAAAVGSAVGTPIALL 371
Db        303 YAERFKKLGYDGDNLLAEYQRGTGTIDASVTAINTALAAIAGGVSAAAAGSVIASPIALL 362

Qy        372 VAGVTGLISGILEASKQAMFESVANRLQGKILEWEKQNGGQNYFDKGYDSRYAAYLANNL 431
Db        363 VSGITGVISTILQYSKQAMFEHVANKIHNKIVEWEKNNHGKNYFENGYDARYLANLQDNM 422

Qy        432 KFLSELNKELEAERVIAITQQRWDNNIGELAGITKLGERIKSGKAYADAFEDGKKVEAGS 491
Db        423 KFLLNLNKELQAERVIAITQQQWDNNIGDLAGISRLGEKVLSGKAYVDAFEEGKHIKADK 482

Qy        492 NITLDAKTGIIDISNSNGKKTQALHFTSPLLTAGTESRERLTNGKYSYINKLKFGRVKNW 551
Db        483 LVQLDSANGIIDVSNSGKAKTQHILFRTPLLTPGTEHRERVQTGKYEYITKLNINRVDSW 542

Qy        552 QVTDGEASSKLDFSKVIQRV-------AETEGTDEIGLIVNAKAGNDDIFVGQGKMNIDG 604
Db        543 KITDGAASSTFDLTNVVQRIGIELDNAGNVTKTKETKIIAKLGEGDDNVFVGSGTTEIDG 602

Qy        605 GDGHDRVFYSKDGGFGNITVDGTSATEAGSYTVNRKVARGDIYHEVVKRQETKVGKRTET 664
Db        603 GEGYDRVHYSR-GNYGALTIDATKETEQGSYTVNRFVETGKALHEVTSTHTALVGNREEK 661

Qy        665 IQYRDYELRKVGYGYQSTDNLKSVEEVIGSQFNDVFKGSKFNDIFHSGEGDDLLDGGAGD 724
Db        662 IEYR-HSNNQHHAGYYTKDTLKAVEEIIGTSHNDIFKGSKFNDAFNGGDGVDTIDGNDGN 720

Qy        725 DRLFGGKGNDRLSGDEGDDLLDGGSGDDVLNGGAGNDVYIFRKGDGNDTLYDGTGNDKLA 784
Db        721 DRLFGGKGDDILDGGNGDDFIDGGKGNDLLHGGKGDDIFVHRKGDGNDIITDSDGNDKLS 780

Qy        785 FADANISDIMIERTKEGIIVKRNDHSGSINIPRWY----ITSNLQNYQSNKTDHKIEQLI 840
Db        781 FSDSNLKDLTFEKVKHNLVI-TNSKKEKVTIQNWFREADFAKEVPNYKATK-DEKIEEII 838

Qy        841 GKDGSYITSDQIDKILQDKKDGTVITSQELKKLADENKSQKLSASDIASSLNKLVGSMAL 900
Db        839 GQNGERITSKQVDDLI--AKGNGKITQDELSKVVDNYELLKHS-KNVTNSLDKLISSVSA 895

Qy        901 FGTANSVSSNALQPITQPTQGI 922
Db        896 FTSSNDSRNVLVAPTSMLDQSL 917
```

```
RESULT 13
AAW79568
ID   AAW79568 standard; Protein; 926 AA.
```

```
XX
AC   AAW79568;
XX
DT   24-DEC-1998  (first entry)
XX
DE   Leukotoxin 352 polypeptide.
XX
KW   Gonadotropin releasing hormone; GnHR; chimera; leukotoxin polypeptide;
KW   multimer; vaccine; tumour; Leukotoxin 352; LKT 352; lktA; plasmid pAA352;
KW   cytotoxic.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Misc_feature    11..924
FT                   /note= "Recombinant leukotoxin peptide"
XX
PN   WO9806848-A1.
XX
PD   19-FEB-1998.
XX
PF   08-AUG-1997;   97WO-CA00559.
XX
PR   09-AUG-1996;   96US-0694865.
XX
PA   (UYSA-) UNIV SASKATCHEWAN.
XX
PI   Manns JG,  Potter AA;
XX
DR   WPI; 1998-159540/14.
DR   N-PSDB; AAV61530.
XX
PT   Chimeric protein of leukotoxin and gonadotropin releasing hormone  -
PT   useful for, e.g. preparation of vaccines for reduction of incidence
PT   of mammary tumours in mammals
XX
PS   Claim 7; Figure 3.1-9; 118pp; English.
XX
CC   The present sequence represents a recombinantly produced or chemically
CC   synthesised leukotoxin 352 (LKT 352)polypeptide, derived from the lktA
CC   gene that is present in the plasmid pAA352.  This gene produces a
CC   truncated protein that has an estimated molecular weight of about 99 kDa
CC   and lacks the cytotoxic portion of the molecule.  Thus this gene has a
CC   higher expression level than that of the full-length molecule.  This can
CC   be used in the construction of a chimeric protein that comprises a
CC   leukotoxin polypeptide, several multimers, and a GnRH sequence.  The
CC   chimeric protein can be used as a vaccine to help reduce the incidence of
CC   mammary tumours in a mammalian individual.
XX
SQ   Sequence  926 AA;

  Query Match            50.1%;  Score 2326.5;  DB 19;  Length 926;
  Best Local Similarity  50.0%;  Pred. No. 1.7e-136;
  Matches  461;  Conservative  174;  Mismatches  262;  Indels   25;  Gaps   12;

Qy         15 LNSTKSGLKNLYLAIPKD--YDPQKGGTLNDFIKAADELGIARLAEEPNHTETAKKSVDT 72
              |:  |:| |  : | ||::  || ::|  | | :|||:||||     || |:  ||: |: |
Db          7 LSFPKTGAKKIILYIPQNYQYDTEQGNGLQDLVKAAEELGIEVQREERNNIATAQTSLGT 66

Qy         73 VNQFLSLTQTGIAISATKLEKFLQKHSTNKLAKGLDSVENIDRKLGKASNVLSTLSSFLG 132
              :   : ||: || :|| :::| |||     |   :  | | |:|  :    ||   |||  : | ||
Db         67 IQTAIGLTERGIVLSAPQIDKLLQK---TKAGQALGSAESIVQNANKAKTVLSGIQSILG 123

Qy        133 TALAGIELDSLIKKGDAAPDALAKASIDLINEIIGNLSQSTQTIEAFSSQLAKLGSTISQ 192
              : |||::||   : : ::     ||||| ::| | :| |:: | :|:: |   |::: || :
Db        124 SVLAGMDLDEAL-QNNSNQHALAKAGLELTNSLIENIANSVKTLDEFGEQISQFGSKLQN 182

Qy        193 AKGFSNIGNKLQNL-NFSKTNLGLEIITGLLSGISAGFALADKNASTGKKVAAGFELSNQ 251
              ||     :|:||:|:       |  |||::|:||||| :|   |||||||| ||| |||||:||
Db        183 IKGLGTLGDKLKNIGGLDKAGLGLDVISGLLSGATAALVLADKNASTAKKVGAGFELANQ 242

Qy        252 VIGNVTKAISSYVLAQRVAAGLSTTGAVAALITSSIMLAISPLAFMNAADKFNHANALDE 311
              |:||:|||:|||:||||||||||:|| ||||| |:: ||||||||    |||||| :|:
Db        243 VVGNITKAVSSYILAQRVAAGLSSTGPVAALIASTVSLAISPLAFAGIADKFNHAKSLES 302

Qy        312 FAKQFRKFGYDGDHLLAEYQRGVGTIEASLTTISTALGAVSAGVSAAAVGSAVGTPIALL 371
              :|::|:| |||||:||||||| |||:||:| |:||| |::  |||||  ||  :  :||||
Db        303 YAERFKKLGYDGDNLLAEYQRGTGTIDASVTAINTALAAIAGGVSAAAAGSVIASPIALL 362

Qy        372 VAGVTGLISGILEASKQAMFESVANRLQGKILEWEKQNGGQNYFDKGYDSRYAAYLANNL 431
              |:|:||:|| ||:  |||||| |||::   ||:||||  |  |:|||:  |||:|| | |  :|:
Db        363 VSGITGVISTILQYSKQAMFEHVANKIHNKIVEWEKNNHGKNYFENGYDARYLANLQDNM 422

Qy        432 KFLSELNKELEAERVIAITQQRWDNNIGELAGITKLGERIKSGKAYADAFEDGKKVEAGS 491
              |||  |||||:||||||||||:||||||:|||||::|||:: |||||  ||||:|| ::|
Db        423 KFLLNLNKELQAERVIAITQQQWDNNIGDLAGISRLGEKVLSGKAYVDAFEEGKHIKADK 482

Qy        492 NITLDAKTGIIDISNSNGKKTQALHFTSPLLTAGTESRERLTNGKYSYINKLKFGRVKNW 551
               :  ||:   ||||:|||     ||| :  | :||||  ||| |||:    ||| || ||   || :|
Db        483 LVQLDSANGIIDVSNSGKAKTQHILFRTPLLTPGTEHRERVQTGKYEYITKLNINRVDSW 542

Qy        552 QVTDGEASSKLDFSKVIQRV-------AETEGTDEIGLIVNAKAGNDDIFVGQGKMNIDG 604
              ::||| |||   |  : |:||:             |  |   :|       |:|::|||  |    |||
Db        543 KITDGAASSTFDLTNVVQRIGIELDNAGNVTKTKETKIIAKLGEGDDNVFVGSGTTEIDG 602

Qy        605 GDGHDRVFYSKDGGFGNITVDGTSATEAGSYTVNRKVARGDIYHEVVKRQETKVGKRTET 664
              |:|:||| ||:  |  :|  :|:|  |    || |||||||  |   |    |||           || |  |
Db        603 GEGYDRVHYSR-GNYGALTIDATKETEQGSYTVNRFVETGKALHEVTSTHTALVGNREEK 661

Qy        665 IQYRDYELRKVGYGYQSTDNLKSVEEVIGSQFNDVFKGSKFNDIFHSGEGDDLLDGGAGD 724
              |:||  :     :     || :  |  ||:|||:||:    ||:||||||||| |: |:| |  :|| |:
Db        662 IEYR-HSNNQHHAGYYTKDTLKAVEEIIGTSHNDIFKGSKFNDAFNGGDGVDTIDGNDGN 720

Qy        725 DRLFGGKGNDRLSGDEGDDLLDGGSGDDVLNGGAGNDVYIFRKGDGNDTLYDGTGNDKLA 784
              ||||||||:|  |  |      |||  :|||  |:|:|:|| |:|:::  ||||||  :  |  |||||:
Db        721 DRLFGGKGDDILDGGNGDDFIDGGKGNDLLHGGKGDDIFVHRKGDGNDIITDSDGNDKLS 780

Qy        785 FADANISDIMIERTKEGIIVKRNDHSGSINIPRWY----ITSNLQNYQSNKTDHKIEQLI 840
              |:|:|: |:   |:  |:  |   :::   |        :.|  |:        : ||:: | | |||::|
Db        781 FSDSNLKDLTFEKVKHNLVI-TNSKKEKVTIQNWFREADFAKEVPNYKATK-DEKIEEII 838

Qy        841 GKDGSYITSDQIDKILQDKKDGTVITSQELKKLADENKSQKLSASDIASSLNKLVGSMAL 900
              |::|  |||  |:|  ::    |    ||   || |:  |    :   | |    :: :||:||: |::
Db        839 GQNGERITSKQVDDLI--AKGNGKITQDELSKVVDNYELLKHS-KNVTNSLDKLISSVSA 895

Qy        901 FGTANSVSSNALQPITQPTQGI 922
              | ::|    :  : | :   | :
Db        896 FTSSNDSRNVLVAPTSMLDQSL 917
```

```
RESULT 14
AAW03942
ID   AAW03942 standard; Protein; 977 AA.
XX
AC   AAW03942;
XX
DT   20-NOV-1996  (first entry)
XX
DE   LKT-GnRH protein fusion from pCB113.
XX
KW   Leukotoxin; LKT; gonadotropin-releasing hormone; GnRH;
KW   fusion protein; immunogen; vaccine; fertility control;
KW   contraceptive; sterilisation.
XX
OS   Chimeric Pasteurella haemolytica A1 strain B122;
OS   Chimeric synthetic.
XX
FH   Key             Location/Qualifiers
FT   Domain          1..929
FT                   /label= LKT
FT   Domain          927..977
FT                   /label= GnRH_repeat_domain
XX
PN   WO9624675-A1.
XX
PD   15-AUG-1996.
```

```
XX
PF   24-JAN-1996;   96WO-CA00049.
XX
PR   10-FEB-1995;   95US-0387156.
XX
PA   (UYSA-) UNIV SASKATCHEWAN.
XX
PI   Manns JG,  Potter AA;
XX
DR   WPI; 1996-384447/38.
DR   N-PSDB; AAT37176.
XX
PT   Gonadotropin-releasing hormone multimer fusion proteins - with
PT   leukotoxin polypeptide for increased immunogenicity, useful in
PT   antifertility vaccine prodn.
XX
PS   Claim 7; Fig 5A-5H; 87pp; English.
XX
CC   A chimeric protein (AAW03942) is composed of a fusion between
CC   a truncated leukotoxin (LKT-352) from Pasteurella haemolytica (see
CC   also AAW03945) and a 4-copy gonadoliberin-releasing hormone (GnRH)
CC   repeat sequence (see also AAW03944).  It is the product of a
CC   chimeric gene (AAT37176) produced by ligating a synthetic sequence
CC   for the 4-copy GnRH into vector pAA352 (ATCC 68283), which carries
CC   the LKT-352 gene.  Recombinant plasmid pCB113 (LKT 352:4 copy
CC   GnRH, ATCC 69749) was obtd.  Escherichia coli transformants
CC   produced the chimeric protein, which is useful as a vaccine for
CC   fertility control, esp. immunological sterilisation of
CC   domestic or farm animals.
XX
SQ   Sequence   977 AA;

  Query Match             50.1%;  Score 2326.5;  DB 17;  Length 977;
  Best Local Similarity   50.0%;  Pred. No. 1.9e-136;
  Matches  461;  Conservative  174;  Mismatches  262;  Indels   25;  Gaps   12;

Qy         15 LNSTKSGLKNLYLAIPKD--YDPQKGGTLNDFIKAADELGIARLAEEPNHTETAKKSVDT 72
Db          7 LSFPKTGAKKIILYIPQNYQYDTEQGNGLQDLVKAAEELGIEVQREERNNIATAQTSLGT 66

Qy         73 VNQFLSLTQTGIAISATKLEKFLQKHSTNKLAKGLDSVENIDRKLGKASNVLSTLSSFLG 132
Db         67 IQTAIGLTERGIVLSAPQIDKLLQK---TKAGQALGSAESIVQNANKAKTVLSGIQSILG 123

Qy        133 TALAGIELDSLIKKGDAAPDALAKASIDLINEIIGNLSQSTQTIEAFSSQLAKLGSTISQ 192
Db        124 SVLAGMDLDEAL-QNNSNQHALAKAGLELTNSLIENIANSVKTLDEFGEQISQFGSKLQN 182

Qy        193 AKGFSNIGNKLQNL-NFSKTNLGLEIITGLLSGISAGFALADKNASTGKKVAAGFELSNQ 251
Db        183 IKGLGTLGDKLKNIGGLDKAGLGLDVISGLLSGATAALVLADKNASTAKKVGAGFELANQ 242

Qy        252 VIGNVTKAISSYVLAQRVAAGLSTTGAVAALITSSIMLAISPLAFMNAADKFNHANALDE 311
Db        243 VVGNITKAVSSYILAQRVAAGLSSTGPVAALIASTVSLAISPLAFAGIADKFNHAKSLES 302

Qy        312 FAKQFRKFGYDGDHLLAEYQRGVGTIEASLTTISTALGAVSAGVSAAAVGSAVGTPIALL 371
Db        303 YAERFKKLGYDGDNLLAEYQRGTGTIDASVTAINTALAAIAGGVSAAAAGSVIASPIALL 362

Qy        372 VAGVTGLISGILEASKQAMFESVANRLQGKILEWEKQNGGQNYFDKGYDSRYAAYLANNL 431
Db        363 VSGITGVISTILQYSKQAMFEHVANKIHNKIVEWEKNNHGKNYFENGYDARYLANLQDNM 422

Qy        432 KFLSELNKELEAERVIAITQQRWDNNIGELAGITKLGERIKSGKAYADAFEDGKKVEAGS 491
Db        423 KFLLNLNKELQAERVIAITQQQWDNNIGDLAGISRLGEKVLSGKAYVDAFEEGKHIKADK 482

Qy        492 NITLDAKTGIIDISNSNGKKTQALHFTSPLLTAGTESRERLTNGKYSYINKLKFGRVKNW 551
Db        483 LVQLDSANGIIDVSNSGKAKTQHILFRTPLLTPGTEHRERVQTGKYEYITKLNINRVDSW 542

Qy        552 QVTDGEASSKLDFSKVIQRV-------AETEGTDEIGLIVNAKAGNDDIFVGQGKMNIDG 604
Db        543 KITDGAASSTFDLTNVVQRIGIELDNAGNVTKTKETKIIAKLGEGDDNVFVGSGTTEIDG 602

Qy        605 GDGHDRVFYSKDGGFGNITVDGTSATEAGSYTVNRKVARGDIYHEVVKRQETKVGKRTET 664
Db        603 GEGYDRVHYSR-GNYGALTIDATKETEQGSYTVNRFVETGKALHEVTSTHTALVGNREEK 661

Qy        665 IQYRDYELRKVGYGYQSTDNLKSVEEVIGSQFNDVFKGSKFNDIFHSGEGDDLLDGGAGD 724
Db        662 IEYR-HSNNQHHAGYYTKDTLKAVEEIIGTSHNDIFKGSKFNDAFNGGDGVDTIDGNDGN 720

Qy        725 DRLFGGKGNDRLSGDEGDDLLDGGSGDDVLNGGAGNDVYIFRKGDGNDTLYDGTGNDKLA 784
Db        721 DRLFGGKGDDILDGGNGDDFIDGGKGNDLLHGGKGDDIFVHRKGDGNDIITDSDGNDKLS 780

Qy        785 FADANISDIMIERTKEGIIVKRNDHSGSINIPRWY----ITSNLQNYQSNKTDHKIEQLI 840
Db        781 FSDSNLKDLTFEKVKHNLVI-TNSKKEKVTIQNWFREADFAKEVPNYKATK-DEKIEEII 838

Qy        841 GKDGSYITSDQIDKILQDKKDGTVITSQELKKLADENKSQKLSASDIASSLNKLVGSMAL 900
Db        839 GQNGERITSKQVDDLI--AKGNGKITQDELSKVVDNYELLKHS-KNVTNSLDKLISSVSA 895

Qy        901 FGTANSVSSNALQPITQPTQGI 922
Db        896 FTSSNDSRNVLVAPTSMLDQSL 917


RESULT 15
AAW79569
ID   AAW79569 standard; Protein; 977 AA.
XX
AC   AAW79569;
XX
DT   24-DEC-1998  (first entry)
XX
DE   LKT-GnRH chimeric protein.
XX
KW   Chimera; pCB113; LKT 352; GnRH; Gonadotropin releasing hormone; multimer;
KW   cytotoxic activity; antigen presentation; immune response; vaccine;
KW   tumour.
XX
OS   Synthetic.
XX
PN   WO9806848-A1.
XX
PD   19-FEB-1998.
XX
PF   08-AUG-1997;   97WO-CA00559.
XX
PR   09-AUG-1996;   96US-0694865.
XX
PA   (UYSA-) UNIV SASKATCHEWAN.
XX
PI   Manns JG,  Potter AA;
XX
DR   WPI; 1998-159540/14.
DR   N-PSDB; AAV61531.
XX
PT   Chimeric protein of leukotoxin and gonadotropin releasing hormone  -
PT   useful for, e.g. preparation of vaccines for reduction of incidence
PT   of mammary tumours in mammals
XX
PS   Disclosure; Figure 5.1-8; 118pp; English.
XX
CC   The present sequence represents the LKT-GnRH chimeric protein from
CC   pCB113.  This plasmid contains the LKT 352 polypeptide (AAW79568) fused
CC   to four copies of the GnRH peptide.  This chimera lacks cytotoxic
CC   activity which enables there to be an increase in antigen presentation
CC   and thus an optimal immune response.  The removal of this region also
CC   enables the truncated LKT to be expressed at much higher levels and
CC   allows the amount of antigen administered to be reduced.  This chimeric
CC   protein comprises a leukotoxin polypeptide, several multimers, and a GnRH
```

```
CC   sequence.  The chimeric protein can be used as a vaccine to help reduce
CC   the incidence of mammary tumours in a mammalian individual.
XX
SQ   Sequence   977 AA;

  Query Match             50.1%;  Score 2326.5;  DB 19;  Length 977;
  Best Local Similarity   50.0%;  Pred. No. 1.9e-136;
  Matches  461;  Conservative  174;  Mismatches  262;  Indels   25;  Gaps   12;

Qy         15 LNSTKSGLKNLYLAIPKD--YDPQKGGTLNDFIKAADELGIARLAEEPNHTETAKKSVDT 72
Db          7 LSFPKTGAKKIILYIPQNYQYDTEQGNGLQDLVKAAEELGIEVQREERNNIATAQTSLGT 66

Qy         73 VNQFLSLTQTGIAISATKLEKFLQKHSTNKLAKGLDSVENIDRKLGKASNVLSTLSSFLG 132
Db         67 IQTAIGLTERGIVLSAPQIDKLLQK---TKAGQALGSAESIVQNANKAKTVLSGIQSILG 123

Qy        133 TALAGIELDSLIKKGDAAPDALAKASIDLINEIIGNLSQSTQTIEAFSSQLAKLGSTISQ 192
Db        124 SVLAGMDLDEAL-QNNSNQHALAKAGLELTNSLIENIANSVKTLDEFGEQISQFGSKLQN 182

Qy        193 AKGFSNIGNKLQNL-NFSKTNLGLEIITGLLSGISAGFALADKNASTGKKVAAGFELSNQ 251
Db        183 IKGLGTLGDKLKNIGGLDKAGLGLDVISGLLSGATAALVLADKNASTAKKVGAGFELANQ 242

Qy        252 VIGNVTKAISSYVLAQRVAAGLSTTGAVAALITSSIMLAISPLAFMNAADKFNHANALDE 311
Db        243 VVGNITKAVSSYILAQRVAAGLSSTGPVAALIASTVSLAISPLAFAGIADKFNHAKSLES 302

Qy        312 FAKQFRKFGYDGDHLLAEYQRGVGTIEASLTTISTALGAVSAGVSAAAVGSAVGTPIALL 371
Db        303 YAERFKKLGYDGDNLLAEYQRGTGTIDASVTAINTALAAIAGGVSAAAAGSVIASPIALL 362

Qy        372 VAGVTGLISGILEASKQAMFESVANRLQGKILEWEKQNGGQNYFDKGYDSRYAAYLANNL 431
Db        363 VSGITGVISTILQYSKQAMFEHVANKIHNKIVEWEKNNHGKNYFENGYDARYLANLQDNM 422

Qy        432 KFLSELNKELEAERVIAITQQRWDNNIGELAGITKLGERIKSGKAYADAFEDGKKVEAGS 491
Db        423 KFLLNLNKELQAERVIAITQQQWDNNIGDLAGISRLGEKVLSGKAYVDAFEEGKHIKADK 482

Qy        492 NITLDAKTGIIDISNSNGKKTQALHFTSPLLTAGTESRERLTNGKYSYINKLKFGRVKNW 551
Db        483 LVQLDSANGIIDVSNSGKAKTQHILFRTPLLTPGTEHRERVQTGKYEYITKLNINRVDSW 542

Qy        552 QVTDGEASSKLDFSKVIQRV-------AETEGTDEIGLIVNAKAGNDDIFVGQGKMNIDG 604
Db        543 KITDGAASSTFDLTNVVQRIGIELDNAGNVTKTKETKIIAKLGEGDDNVFVGSGTTEIDG 602

Qy        605 GDGHDRVFYSKDGGFGNITVDGTSATEAGSYTVNRKVARGDIYHEVVKRQETKVGKRTET 664
Db        603 GEGYDRVHYSR-GNYGALTIDATKETEQGSYTVNRFVETGKALHEVTSTHTALVGNREEK 661

Qy        665 IQYRDYELRKVGYGYQSTDNLKSVEEVIGSQFNDVFKGSKFNDIFHSGEGDDLLDGGAGD 724
Db        662 IEYR-HSNNQHHAGYYTKDTLKAVEEIIGTSHNDIFKGSKFNDAFNGGDGVDTIDGNDGN 720

Qy        725 DRLFGGKGNDRLSGDEGDDLLDGGSGDDVLNGGAGNDVYIFRKGDGNDTLYDGTGNDKLA 784
Db        721 DRLFGGKGDDILDGGNGDDFIDGGKGNDLLHGGKGDDIFVHRKGDGNDIITDSDGNDKLS 780

Qy        785 FADANISDIMIERTKEGIIVKRNDHSGSINIPRWY----ITSNLQNYQSNKTDHKIEQLI 840
Db        781 FSDSNLKDLTFEKVKHNLVI-TNSKKEKVTIQNWFREADFAKEVPNYKATK-DEKIEEII 838

Qy        841 GKDGSYITSDQIDKILQDKKDGTVITSQELKKLADENKSQKLSASDIASSLNKLVGSMAL 900
Db        839 GQNGERITSKQVDDLI--AKGNGKITQDELSKVVDNYELLKHS-KNVTNSLDKLISSVSA 895

Qy        901 FGTANSVSSNALQPITQPTQGI 922
Db        896 FTSSNDSRNVLVAPTSMLDQSL 917
```

```
RESULT 16
AAR52748
ID   AAR52748 standard; Protein; 1069 AA.
XX
AC   AAR52748;
XX
DT   25-MAR-2003  (updated)
DT   01-JUL-1994  (first entry)
XX
DE   Bovine IFNgamma/LKT chimeric protein encoded by plasmid pAA497.
XX
KW   Bovine; interleukin-2; IL2; P. haemolytica; leukotoxin; LTK; IFN;
KW   ltkA; chromosome walking; fusion protein; vaccine; interferon; gamma;
KW   monoclonal; polyclonal; antibody.
XX
OS   Pasteurella heamolytica - chimera.
OS   Bos taurus - chimera.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1..926
FT                   /note= "Recombinant leukotoxin peptide [split]"
FT   Peptide         927..1069
FT                   /note= "Bovine IFNgamma"
XX
PN   US5273889-A.
XX
PD   28-DEC-1993.
XX
PF   16-OCT-1991;   91US-0777715.
XX
PR   22-AUG-1990;   90US-0571301.
PR   16-OCT-1991;   91US-0777715.
XX
PA   (CIBA ) CIBA GEIGY CANADA LTD.
PA   (UYSA-) UNIV SASKATCHEWAN.
XX
PI   Campos M,  Hughes HPA,  Potter A;
XX
DR   WPI; 1994-006687/01.
DR   N-PSDB; AAQ54213.
XX
PT   Immunogenic fusion proteins of gamma-interferon and immunogenic
PT   leukotoxin - used in vaccines and to raise monoclonal and polyclonal
PT   antibodies
XX
PS   Disclosure; Fig 7; 56pp; English.
XX
CC   This sequence represents a fusion between bovine gamma interferon
CC   (IFNgamma) and Pasteurella haemolytica leukotoxin (LTK).  The
CC   leukotoxin gene, ltkA, was isolated from a gene library of P.
CC   haemolytica by chromosome walking.  Immunogenic fusion proteins
CC   such as this can be used in vaccine compositions.  It can also be
CC   used to raise mono- and polyclonal antibodies.
CC   (Updated on 25-MAR-2003 to correct PF field.)
XX
SQ   Sequence   1069 AA;

  Query Match             50.1%;  Score 2326.5;  DB 15;  Length 1069;
  Best Local Similarity   50.0%;  Pred. No. 2.1e-136;
  Matches  461;  Conservative  174;  Mismatches  262;  Indels   25;  Gaps   12;

Qy         15 LNSTKSGLKNLYLAIPKD--YDPQKGGTLNDFIKAADELGIARLAEEPNHTETAKKSVDT 72
Db          7 LSFPKTGAKKIILYIPQNYQYDTEQGNGLQDLVKAAEELGIEVQREERNNIATAQTSLGT 66

Qy         73 VNQFLSLTQTGIAISATKLEKFLQKHSTNKLAKGLDSVENIDRKLGKASNVLSTLSSFLG 132
Db         67 IQTAIGLTERGIVLSAPQIDKLLQK---TKAGQALGSAESIVQNANKAKTVLSGIQSILG 123

Qy        133 TALAGIELDSLIKKGDAAPDALAKASIDLINEIIGNLSQSTQTIEAFSSQLAKLGSTISQ 192
Db        124 SVLAGMDLDEAL-QNNSNQHALAKAGLELTNSLIENIANSVKTLDEFGEQISQFGSKLQN 182
```

```
Qy        193 AKGFSNIGNKLQNL-NFSKTNLGLEIITGLLSGISAGFALADKNASTGKKVAAGFELSNQ 251
              ||   :|:||:|:      |  |||::|:||||| :|   ||||||| ||| |||||:||
Db        183 IKGLGTLGDKLKNIGGLDKAGLGLDVISGLLSGATAALVLADKNASTAKKVGAGFELANQ 242

Qy        252 VIGNVTKAISSYVLAQRVAAGLSTTGAVAALITSSIMLAISPLAFMNAADKFNHANALDE 311
              |:||:|||:|||:|||||||||||:|| ||||| |:: |||||||    |||||| :|:
Db        243 VVGNITKAVSSYILAQRVAAGLSSTGPVAALIASTVSLAISPLAFAGIADKFNHAKSLES 302

Qy        312 FAKQFRKFGYDGDHLLAEYQRGVGTIEASLTTISTALGAVSAGVSAAAVGSAVGTPIALL 371
               :|::|:| |||||:||||||| |||:||:| |:||| |:: ||||| || : :||||
Db        303 YAERFKKLGYDGDNLLAEYQRGTGTIDASVTAINTALAAIAGGVSAAAAGSVIASPIALL 362

Qy        372 VAGVTGLISGILEASKQAMFESVANRLQGKILEWEKQNGGQNYFDKGYDSRYAAYLANNL 431
              |:|:||:|| ||: |||||| |||::  ||:|||| | |:|||: |||:|| | | :|:
Db        363 VSGITGVISTILQYSKQAMFEHVANKIHNKIVEWEKNNHGKNYFENGYDARYLANLQDNM 422

Qy        432 KFLSELNKELEAERVIAITQQRWDNNIGELAGITKLGERIKSGKAYADAFEDGKKVEAGS 491
              |||  |||||:|||||||||:||||||:||||::|||:: ||||| ||||:|| ::|
Db        423 KFLLNLNKELQAERVIAITQQQWDNNIGDLAGISRLGEKVLSGKAYVDAFEEGKHIKADK 482

Qy        492 NITLDAKTGIIDISNSNGKKTQALHFTSPLLTAGTESRERLTNGKYSYINKLKFGRVKNW 551
               : ||:  ||||:|||   |||  :  | :||||  ||| |||:  ||| || ||   || :|
Db        483 LVQLDSANGIIDVSNSGKAKTQHILFRTPLLTPGTEHRERVQTGKYEYITKLNINRVDSW 542

Qy        552 QVTDGEASSKLDFSKVIQRV-------AETEGTDEIGLIVNAKAGNDDIFVGQGKMNIDG 604
              ::||| |||   | : |:||:                  | |  :|     |:|::|||  |  |||
Db        543 KITDGAASSTFDLTNVVQRIGIELDNAGNVTKTKETKIIAKLGEGDDNVFVGSGTTEIDG 602

Qy        605 GDGHDRVFYSKDGGFGNITVDGTSATEAGSYTVNRKVARGDIYHEVVKRQETKVGKRTET 664
              |:|:||| ||:  | :| :|:| |  || |||||||  |   |    |||           || | |
Db        603 GEGYDRVHYSR-GNYGALTIDATKETEQGSYTVNRFVETGKALHEVTSTHTALVGNREEK 661

Qy        665 IQYRDYELRKVGYGYQSTDNLKSVEEVIGSQFNDVFKGSKFNDIFHSGEGDDLLDGGAGD 724
              |:|| :      ||  :  | ||:|||:||:  ||:||||||||| |:  |:| | :|| |:
Db        662 IEYR-HSNNQHHAGYYTKDTLKAVEEIIGTSHNDIFKGSKFNDAFNGGDGVDTIDGNDGN 720

Qy        725 DRLFGGKGNDRLSGDEGDDLLDGGSGDDVLNGGAGNDVYIFRKGDGNDTLYDGTGNDKLA 784
              ||||||||:| |  ||| :||| |:|:|:||  |:|:::| ||||||  : |  ||||:
Db        721 DRLFGGKGDDILDGGNGDDFIDGGKGNDLLHGGKGDDIFVHRKGDGNDIITDSDGNDKLS 780

Qy        785 FADANISDIMIERTKEGIIVKRNDHSGSINIPRWY----ITSNLQNYQSNKTDHKIEQLI 840
              |:|:|: |:  |: |  ::: |       : |  |:       : ||:: | | |||::|
Db        781 FSDSNLKDLTFEKVKHNLVI-TNSKKEKVTIQNWFREADFAKEVPNYKATK-DEKIEEII 838

Qy        841 GKDGSYITSDQIDKILQDKKDGTVITSQELKKLADENKSQKLSASDIASSLNKLVGSMAL 900
              |::|  ||| |:| ::    |      ||  || |: |   :  | |    :: :|:||: |::
Db        839 GQNGERITSKQVDDLI--AKGNGKITQDELSKVVDNYELLKHS-KNVTNSLDKLISSVSA 895

Qy        901 FGTANSVSSNALQPITQPTQGI 922
              | ::|   :  : | :   | :
Db        896 FTSSNDSRNVLVAPTSMLDQSL 917
```

```
RESULT 17
AAW13867
ID   AAW13867 standard; Protein; 1069 AA.
XX
AC   AAW13867;
XX
DT   25-MAR-2003  (updated)
DT   12-MAY-1997  (first entry)
XX
DE   Chimeric protein #2.
XX
KW   RTX cytotoxin; cytokine; immunogen; chimeric protein; cytokine; vaccine;
KW   interleukin-2; IL-2; gamma interferon; gamma IFN; leukotoxin; pneumonia;
KW   Pasteurella haemolytica; LKT352; respiratory disease; shipping fever;
KW   fibrinous pneumonia; cattle; therapy.
XX
OS   Synthetic.
XX
PN   US5594107-A.
XX
PD   14-JAN-1997.
XX
PF   20-DEC-1993;   93US-0170126.
XX
PR   20-DEC-1993;   93US-0170126.
PR   22-AUG-1990;   90US-0571301.
PR   16-OCT-1991;   91US-0777715.
XX
PA   (CIBA ) CIBA GEIGY CANADA LTD.
PA   (UYSA-) UNIV SASKATCHEWAN.
XX
PI   Campos M,  Hughes HPA,  Potter A;
XX
DR   WPI; 1997-099529/09.
DR   N-PSDB; AAT60033.
XX
PT   Immunogenic chimeric proteins comprising cytokine linked to RTX
PT   toxin - useful in vaccines, esp. against shipping fever in cattle
XX
PS   Claim 13; Column 37-46; 56pp; English.
XX
CC   AAW13866 and AAW13867 represent immunogenic chimeric proteins of the
CC   invention.  This sequence represents a chimeric protein containing the
CC   bovine gamma interferon (gamma IFN) sequence and a leukotoxin sequence.
CC   The chimeric proteins of the invention comprise a cytokine, selected
CC   from interleukin-2 (IL-2) and gamma IFN, linked to at least one RTX
CC   toxin epitope (preferably the sequence shown in AAW13865).  The RTX toxin
CC   used to provide the epitope sequence is preferably a leukotoxin,
CC   especially the full-length Pasteurella haemolytica leukotoxin.
CC   Alternatively, the leukotoxin is a truncated leukotoxin lacking
CC   leukotoxic activity, especially LKT352.  The chimeric proteins can be
CC   used for the production if vaccines against respiratory diseases such as
CC   pneumonia, particularly fibrinous pneumonia caused by P.haemolytica,
CC   including shipping fever in cattle.
CC   (Updated on 25-MAR-2003 to correct PF field.)
XX
SQ   Sequence   1069 AA;

  Query Match             50.1%;  Score 2326.5;  DB 18;  Length 1069;
  Best Local Similarity   50.0%;  Pred. No. 2.1e-136;
  Matches  461;  Conservative  174;  Mismatches  262;  Indels   25;  Gaps   12;

Qy         15 LNSTKSGLKNLYLAIPKD--YDPQKGGTLNDFIKAADELGIARLAEEPNHTETAKKSVDT 72
              |:   |:|  | :  | ||::     |  | ::|     | | :|||:||||      || |:  ||: |: |
Db          7 LSFPKTGAKKIILYIPQNYQYDTEQGNGLQDLVKAAEELGIEVQREERNNIATAQTSLGT 66

Qy         73 VNQFLSLTQTGIAISATKLEKFLQKHSTNKLAKGLDSVENIDRKLGKASNVLSTLSSFLG 132
              :   :  ||: || :|| :::| |||       |   :  |  | |:|  :     ||  |||  : | ||
Db         67 IQTAIGLTERGIVLSAPQIDKLLQK---TKAGQALGSAESIVQNANKAKTVLSGIQSILG 123

Qy        133 TALAGIELDSLIKKGDAAPDALAKASIDLINEIIGNLSQSTQTIEAFSSQLAKLGSTISQ 192
              : |||::||   : : ::     |||||  ::|  |  :|  |:: | :|:: |   |::: || :
Db        124 SVLAGMDLDEAL-QNNSNQHALAKAGLELTNSLIENIANSVKTLDEFGEQISQFGSKLQN 182

Qy        193 AKGFSNIGNKLQNL-NFSKTNLGLEIITGLLSGISAGFALADKNASTGKKVAAGFELSNQ 251
              ||   :|:||:|:      |  |||::|:||||| :|   ||||||| ||| |||||:||
Db        183 IKGLGTLGDKLKNIGGLDKAGLGLDVISGLLSGATAALVLADKNASTAKKVGAGFELANQ 242

Qy        252 VIGNVTKAISSYVLAQRVAAGLSTTGAVAALITSSIMLAISPLAFMNAADKFNHANALDE 311
              |:||:|||:|||:|||||||||||:|| ||||| |:: |||||||    |||||| :|:
Db        243 VVGNITKAVSSYILAQRVAAGLSSTGPVAALIASTVSLAISPLAFAGIADKFNHAKSLES 302

Qy        312 FAKQFRKFGYDGDHLLAEYQRGVGTIEASLTTISTALGAVSAGVSAAAVGSAVGTPIALL 371
               :|::|:| |||||:||||||| |||:||:| |:||| |:: ||||| || : :||||
Db        303 YAERFKKLGYDGDNLLAEYQRGTGTIDASVTAINTALAAIAGGVSAAAAGSVIASPIALL 362

Qy        372 VAGVTGLISGILEASKQAMFESVANRLQGKILEWEKQNGGQNYFDKGYDSRYAAYLANNL 431
              |:|:||:|| ||: |||||| |||::  ||:|||| | |:|||: |||:|| | | :|:
Db        363 VSGITGVISTILQYSKQAMFEHVANKIHNKIVEWEKNNHGKNYFENGYDARYLANLQDNM 422

Qy        432 KFLSELNKELEAERVIAITQQRWDNNIGELAGITKLGERIKSGKAYADAFEDGKKVEAGS 491
              |||  |||||:||||||||||:||||||:||||::|||:: ||||| ||||:|| ::|
```

```
Db        423 KFLLNLNKELQAERVIAITQQQWDNNIGDLAGISRLGEKVLSGKAYVDAFEEGKHIKADK 482

Qy        492 NITLDAKTGIIDISNSNGKKTQALHFTSPLLTAGTESRERLTNGKYSYINKLKFGRVKNW 551
              : ||:  ||||:|||   |||  : | :|||| ||| |||:  ||| || ||    || :|
Db        483 LVQLDSANGIIDVSNSGKAKTQHILFRTPLLTPGTEHRERVQTGKYEYITKLNINRVDSW 542

Qy        552 QVTDGEASSKLDFSKVIQRV-------AETEGTDEIGLIVNAKAGNDDIFVGQGKMNIDG 604
              ::||| |||  | : |:||:               | |  :|      |:|::||| |   |||
Db        543 KITDGAASSTFDLTNVVQRIGIELDNAGNVTKTKETKIIAKLGEGDDNVFVGSGTTEIDG 602

Qy        605 GDGHDRVFYSKDGGFGNITVDGTSATEAGSYTVNRKVARGDIYHEVVKRQETKVGKRTET 664
              |:|:||| ||: | :| :|:| |  || ||||||| |   |  |||       || | |
Db        603 GEGYDRVHYSR-GNYGALTIDATKETEQGSYTVNRFVETGKALHEVTSTHTALVGNREEK 661

Qy        665 IQYRDYELRKVGYGYQSTDNLKSVEEVIGSQFNDVFKGSKFNDIFHSGEGDDLLDGGAGD 724
              |:||  :   :    || :  | ||:|||:||:  ||:||||||||| |: |:| | :|| |:
Db        662 IEYR-HSNNQHHAGYYTKDTLKAVEEIIGTSHNDIFKGSKFNDAFNGGDGVDTIDGNDGN 720

Qy        725 DRLFGGKGNDRLSGDEGDDLLDGGSGDDVLNGGAGNDVYIFRKGDGNDTLYDGTGNDKLA 784
              ||||||||:| | |  |||  :||| |:|:|:|| |:|::: |||||| :  |   |||||:
Db        721 DRLFGGKGDDILDGGNGDDFIDGGKGNDLLHGGKGDDIFVHRKGDGNDIITDSDGNDKLS 780

Qy        785 FADANISDIMIERTKEGIIVKRNDHSGSINIPRWY----ITSNLQNYQSNKTDHKIEQLI 840
              |:|:|: |:  |: |  :::  |       :  |  |:          :  ||::  | | |||::|
Db        781 FSDSNLKDLTFEKVKHNLVI-TNSKKEKVTIQNWFREADFAKEVPNYKATK-DEKIEEII 838

Qy        841 GKDGSYITSDQIDKILQDKKDGTVITSQELKKLADENKSQKLSASDIASSLNKLVGSMAL 900
              |::|  |||  |:|  ::   |     ||   || |: |   :   |  |  :: :||:||:  |::
Db        839 GQNGERITSKQVDDLI--AKGNGKITQDELSKVVDNYELLKHS-KNVTNSLDKLISSVSA 895

Qy        901 FGTANSVSSNALQPITQPTQGI 922
              | ::|   :  : |  :  | :
Db        896 FTSSNDSRNVLVAPTSMLDQSL 917
```

```
RESULT 18
AAB21074
ID   AAB21074 standard; Protein; 1069 AA.
XX
AC   AAB21074;
XX
DT   19-DEC-2000  (first entry)
XX
DE   Bovine gamma-IFN/Pasteurella haemolytica leukotoxin fusion gene.
XX
KW   Bovine gamma-IFN; gamma-interferon; leukotoxin; LKT; respiratory disease;
KW   pneumonia; shipping fever; cattle; livestock; anti-Pasteurella vaccine;
KW   immunogen.
XX
OS   Chimeric - Bos taurus.
OS   Chimeric - Pasteurella haemolytica.
XX
PN   US6096320-A.
XX
PD   01-AUG-2000.
XX
PF   20-OCT-1997;   97US-0954418.
XX
PR   20-DEC-1993;   93US-0170126.
PR   22-JUL-1996;   96US-0681479.
PR   22-AUG-1990;   90US-0571301.
PR   16-OCT-1991;   91US-0777715.
XX
PA   (UYSA-) UNIV SASKATCHEWAN.
PA   (CIBA ) CIBA GEIGY CANADA LTD.
XX
PI   Campos M,  Hughes HPA,  Potter A;
XX
DR   WPI; 2000-531543/48.
DR   N-PSDB; AAA72484.
XX
PT   Vaccine for stimulating immunity against pneumonia comprises chimeric
PT   protein comprising gamma-interferon and leukotoxin derived from
PT   Pasteurella haemolytica  -
XX
PS   Claim 8; Column 45-50; 56pp; English.
XX
CC   This sequence a fusion protein comprising bovine gamma-interferon
CC   (gamma-IFN) and Pasteurella haemolytica leukotoxin (LKT). The
CC   fusion protein is immunogenic, and may be used in an anti-
CC   Pasteurella vaccine composition. Pasteurella species, especially
CC   Pasteurella haemolytica, are responsible for respiratory diseases
CC   in a range of agricultural animals, most particularly cattle, but also
CC   sheep, pigs, horses and fowl. Shipping fever is the most economically
CC   important respiratory disease associated with Pasteurella species,
CC   affecting 15-30% of exposed cattle and resulting in a 2-5% mortality
CC   rate in the exposed population. The vaccine composition of the
CC   invention is is useful for preventing or ameliorating respiratory
CC   diseases such as pneumonia, particularly shipping fever pneumonia, in
CC   livestock.
XX
SQ   Sequence   1069 AA;

  Query Match             50.1%;  Score 2326.5;  DB 21;  Length 1069;
  Best Local Similarity   50.0%;  Pred. No. 2.1e-136;
  Matches  461;  Conservative  174;  Mismatches  262;  Indels   25;  Gaps   12;

Qy         15 LNSTKSGLKNLYLAIPKD--YDPQKGGTLNDFIKAADELGIARLAEEPNHTETAKKSVDT 72
              |:  |:| | : | ||::  || :   | | :|||:||||       || |:  ||: |: |
Db          7 LSFPKTGAKKIILYIPQNYQYDTEQGNGLQDLVKAAEELGIEVQREERNNIATAQTSLGT 66

Qy         73 VNQFLSLTQTGIAISATKLEKFLQKHSTNKLAKGLDSVENIDRKLGKASNVLSTLSSFLG 132
              :   : ||: || :|| :::| |||      |  : | | |:| :      ||  |||  : | ||
Db         67 IQTAIGLTERGIVLSAPQIDKLLQK---TKAGQALGSAESIVQNANKAKTVLSGIQSILG 123

Qy        133 TALAGIELDSLIKKGDAAPDALAKASIDLINEIIGNLSQSTQTIEAFSSQLAKLGSTISQ 192
              : |||::||  : : ::  |||| ::| | :| |:: | :|::: |  |::: || :
Db        124 SVLAGMDLDEAL-QNNSNQHALAKAGLELTNSLIENIANSVKTLDEFGEQISQFGSKLQN 182

Qy        193 AKGFSNIGNKLQNL-NFSKTNLGLEIITGLLSGISAGFALADKNASTGKKVAAGFELSNQ 251
              ||   :|:||:|:  |   |||::|:||||  :   ||||||||| ||| ||||:||
Db        183 IKGLGTLGDKLKNIGGLDKAGLGLDVISGLLSGATAALVLADKNASTAKKVGAGFELANQ 242

Qy        252 VIGNVTKAISSYVLAQRVAAGLSTTGAVAALITSSIMLAISPLAFMNAADKFNHANALDE 311
              |:||:|||:|||:|||||||||||:|| ||||| |:: ||||||||   |||||||  :|:
Db        243 VVGNITKAVSSYILAQRVAAGLSSTGPVAALIASTVSLAISPLAFAGIADKFNHAKSLES 302

Qy        312 FAKQFRKFGYDGDHLLAEYQRGVGTIEASLTTISTALGAVSAGVSAAAVGSAVGTPIALL 371
              :|::|:| |||||:|||||||| |||:|:|:| |:|||  |:: |||||  ||  :  :|||||
Db        303 YAERFKKLGYDGDNLLAEYQRGTGTIDASVTAINTALAAIAGGVSAAAAGSVIASPIALL 362

Qy        372 VAGVTGLISGILEASKQAMFESVANRLQGKILEWEKQNGGQNYFDKGYDSRYAAYLANNL 431
              |:|:||:|| ||:  |||||| |||::  ||:||||  | |:|||: |||:|| |  | :|:
Db        363 VSGITGVISTILQYSKQAMFEHVANKIHNKIVEWEKNNHGKNYFENGYDARYLANLQDNM 422

Qy        432 KFLSELNKELEAERVIAITQQRWDNNIGELAGITKLGERIKSGKAYADAFEDGKKVEAGS 491
              |||  |||||:||||||||||:|||||:||||::|||:: ||||| |||| ||:|| ::|
Db        423 KFLLNLNKELQAERVIAITQQQWDNNIGDLAGISRLGEKVLSGKAYVDAFEEGKHIKADK 482

Qy        492 NITLDAKTGIIDISNSNGKKTQALHFTSPLLTAGTESRERLTNGKYSYINKLKFGRVKNW 551
              : ||:  ||||:|||   |||  : | :|||| ||| |||:  ||| || ||    || :|
Db        483 LVQLDSANGIIDVSNSGKAKTQHILFRTPLLTPGTEHRERVQTGKYEYITKLNINRVDSW 542

Qy        552 QVTDGEASSKLDFSKVIQRV-------AETEGTDEIGLIVNAKAGNDDIFVGQGKMNIDG 604
              ::||| |||  | : |:||:               | |  :|      |:|::||| |   |||
Db        543 KITDGAASSTFDLTNVVQRIGIELDNAGNVTKTKETKIIAKLGEGDDNVFVGSGTTEIDG 602

Qy        605 GDGHDRVFYSKDGGFGNITVDGTSATEAGSYTVNRKVARGDIYHEVVKRQETKVGKRTET 664
              |:|:||| ||: | :| :|:| |  || ||||||| |   |  |||       || | |
Db        603 GEGYDRVHYSR-GNYGALTIDATKETEQGSYTVNRFVETGKALHEVTSTHTALVGNREEK 661

Qy        665 IQYRDYELRKVGYGYQSTDNLKSVEEVIGSQFNDVFKGSKFNDIFHSGEGDDLLDGGAGD 724
              |:||  :   :    || :  | ||:|||:||:  ||:||||||||| |: |:| | :|| |:
Db        662 IEYR-HSNNQHHAGYYTKDTLKAVEEIIGTSHNDIFKGSKFNDAFNGGDGVDTIDGNDGN 720

Qy        725 DRLFGGKGNDRLSGDEGDDLLDGGSGDDVLNGGAGNDVYIFRKGDGNDTLYDGTGNDKLA 784
```

```
               ||||||||:| |  |  |||  :|||  |:|:|:||  |:|:::  ||||||| : |  |||||:
Db       721 DRLFGGKGDDILDGGNGDDFIDGGKGNDLLHGGKGDDIFVHRKGDGNDIITDSDGNDKLS 780

Qy       785 FADANISDIMIERTKEGIIVKRNDHSGSINIPRWY----ITSNLQNYQSNKTDHKIEQLI 840
             |:|:|: |: |: |  ::: |      : |  |:         : ||:: | | |||::|
Db       781 FSDSNLKDLTFEKVKHNLVI-TNSKKEKVTIQNWFREADFAKEVPNYKATK-DEKIEEII 838

Qy       841 GKDGSYITSDQIDKILQDKKDGTVITSQELKKLADENKSQKLSASDIASSLNKLVGSMAL 900
             |::|  ||| |:| ::    |    ||  || |: |   :  | |   :: :||:||: |::
Db       839 GQNGERITSKQVDDLI--AKGNGKITQDELSKVVDNYELLKHS-KNVTNSLDKLISSVSA 895

Qy       901 FGTANSVSSNALQPITQPTQGI 922
             | ::|    :  : |  :   | :
Db       896 FTSSNDSRNVLVAPTSMLDQSL 917
```

RESULT 19
AAR07167
ID AAR07167 standard; protein; 953 AA.
XX
AC AAR07167;
XX
DT 24-JAN-1991 (first entry)
XX
DE 105kD PTX protein of P.haemolytica.
XX
KW Vaccine; shipping fever.
XX
OS Pasteurella haemolytica.
XX
PN US4957739-A.
XX
PD 18-SEP-1990.
XX
PF 13-AUG-1987; 87US-0085430.
XX
PR 13-AUG-1987; 87US-0085430.
XX
PA (TEXA) UNIV OF TEXAS SYSTE.
XX
PI Berget P, Engler M, Highlander S, Weinstock G;
XX
DR WPI; 1990-304558/40.
DR N-PSDB; AAQ06074.
XX
PT Vaccine against shipping fever in cattle - contains purified
PT Pasteurella haemolytica antigen of molecular wt. 105 kD
XX
PS Disclosure; Fig 9a-h; 35pp; English.
XX
CC Gene product is antigenic to P.haemolytica, and may be used as a
CC vaccine for immunisation against shipping fever. Abs raised to the
CC antigen may be used in passive immunisation and diagnosis.
XX
SQ Sequence 953 AA;

```
  Query Match             50.0%;  Score 2325;  DB 11;  Length 953;
  Best Local Similarity   49.9%;  Pred. No. 2.2e-136;
  Matches  464;  Conservative  175;  Mismatches  264;  Indels   26;  Gaps   13;

Qy         8 KSNIQAGLNSTKSGLKNLYLAIPKD--YDPQKGGTLNDFIKAADELGIARLAEEPNHTET 65
             :|   ||| :| |:| |  : | ||::   || ::|  | | :|||:||||    || |:  |
Db        28 QSLTQAG-SSLKTGAKKIILYIPQNYQYDTEQGNGLQDLVKAAEELGIEVQREERNNIAT 86

Qy        66 AKKSVDTVNQFLSLTQTGIAISATKLEKFLQKHSTNKLAKGLDSVENIDRKLGKASNVLS 125
             |: |: |:    : ||: || :|| :::| |||      |  : | | |:| :    ||  |||
Db        87 AQTSLGTIQTAIGLTERGIVLSAPQIDKLLQK---TKAGQALGSAESIVQNANKAKTVLS 143

Qy       126 TLSSFLGTALAGIELDSLIKKGDAAPDALAKASIDLINEIIGNLSQSTQTIEAFSSQLAK 185
              : | ||: |||::||  : : ::   ||||| ::| | :| |:: | :|:: |  |:::
Db       144 GIQSILGSVLAGMDLDEAL-QNNSNQHALAKAGLELTNSLIENIANSVKTLDEFGEQISQ 202

Qy       186 LGSTISQAKGFSNIGNKLQNL-NFSKTNLGLEIITGLLSGISAGFALADKNASTGKKVAA 244
             || :   ||   :|:||:      |  |||::|:||||  :|   ||||||| ||| |
Db       203 FGSKLQNIKGLGTLGDKLKNIGGLDKAGLGLDVISGLLSGATAALVLADKNASTAKKVGA 262

Qy       245 GFELSNQVIGNVTKAISSYVLAQRVAAGLSTTGAVAALITSSIMLAISPLAFMNAADKFN 304
             ||||:|||:||:|||:|||:||||||||||:|| |||||  |:: ||||||||    |||||
Db       263 GFELANQVVGNITKAVSSYILAQRVAAGLSSTGPVAALIASTVSLAISPLAFAGIADKFN 322

Qy       305 HANALDEFAKQFRKFGYDGDHLLAEYQRGVGTIEASLTTISTALGAVSAGVSAAAVGSAV 364
             || :|: :|:|:|  ||||:||||||||  |||:|:|: |:|||  |:: ||||| || :
Db       323 HAKSLESYAERFKKLGYDGDNLLAEYQRGTGTIDASVTAINTALAAIAGGVSAAAAGSVI 382

Qy       365 GTPIALLVAGVTGLISGILEASKQAMFESVANRLQGKILEWEKQNGGQNYFDKGYDSRYA 424
             :||||||:|:||:|| ||: ||||||| |||::  ||:|||| |:|:: |:|||: |||:||
Db       383 ASPIALLVSGITGVISTILQYSKQAMFEHVANKIHNKIVEWEKNNHGKNYFENGYDARYL 442

Qy       425 AYLANNLKFLSELNKELEAERVIAITQQRWDNNIGELAGITKLGERIKSGKAYADAFEDG 484
             | | :|:|||  |||||:||||||||||:||||||:||||:|||::|||:: |||||  ||||:|
Db       443 ANLQDNMKFLLNLNKELQAERVIAITQQQWDNNIGDLAGISRLGEKVLSGKAYVDAFEEG 502

Qy       485 KKVEAGSNITLDAKTGIIDISNSNGKKTQALHFTSPLLTAGTESRERLTNGKYSYINKLK 544
             | ::|   : ||: |||:|||   ||| : | :|||| ||| |||: ||| || |
Db       503 KHIKADKLVQLDSANGIIDVSNSGKAKTQHILFRTPLLTPGTEHRERVQTGKYEYITKLN 562

Qy       545 FGRVKNWQVTDGEASSKLDFSKVIQRV-------AETEGTDEIGLIVNAKAGNDDIFVGQ 597
             || :|::||| |||  | : |:||:                     | |   :|   |:|::|||
Db       563 INRVDSWKITDGAASSTFDLTNVVQRIGIELDNAGNVTKTKETKIIAKLGEGDDNVFVGS 622

Qy       598 GKMNIDGGDGHDRVFYSKDGGFGNITVDGTSATEAGSYTVNRKVARGDIYHEVVKRQETK 657
             |   ||||:|:|||  ||:   | :| :|:|  | |||||||   |  | | |||
Db       623 GTTEIDGGEGYDRVHYSR-GNYGALTIDATKETEQGSYTVNRFVETGKALHEVTSTHTAL 681

Qy       658 VGKRTETIQYRDYELRKVGYGYQSTDNLKSVEEVIGSQFNDVFKGSKFNDIFHSGEGDDL 717
             || | |  |:|| :   :   || :  | ||:|||:||: ||:|||||||| |: |:| |
Db       682 VGNREEKIEYR-HSNNQHHAGYYTKDTLKAVEEIIGTSHNDIFKGSKFNDAFNGGDGVDT 740

Qy       718 LDGGAGDDRLFGGKGNDRLSGDEGDDLLDGGSGDDVLNGGAGNDVYIFRKGDGNDTLYDG 777
             : |  |:|||||||||:| |  | |||  :|||  |:|:|:|||  |:|:::  ||||||| : |
Db       741 IYGNDGNDRLFGGKGDDILDGGNGDDFIDGGKGNDLLHGGKGDDIFVHRKGDGNDIITDS 800

Qy       778 TGNDKLAFADANISDIMIERTKEGIIVKRNDHSGSINIPRWY----ITSNLQNYQSNKTD 833
              |||||:|:|:|:|: |:  |:  |: |  :::  |      |  |:             : ||:: | |
Db       801 DGNDKLSFSDSNLKDLTFEKVKHNLVI-TNSKKEKVTIQNWFREADFAKEVPNYKATK-D 858

Qy       834 HKIEQLIGKDGSYITSDQIDKILQDKKDGTVITSQELKKLADENKSQKLSASDIASSLNK 893
              |||::||::|  ||| |:| ::    ||  || |: |  :  | |  :: :||:|
Db       859 EKIEEIIGQNGERITSKQVDDLI--AKGNGKITQDELSKVVDNYELLKHS-KNVTNSLDK 915

Qy       894 LVGSMALFGTANSVSSNALQPITQPTQGI 922
             |: |:: | ::|    :  : | :   | :
Db       916 LISSVSAFTSSNDSRNVLVAPTSMLDQSL 944
```

RESULT 20
AAR60072
ID AAR60072 standard; Protein; 953 AA.
XX
AC AAR60072;
XX
DT 25-MAR-2003 (updated)
DT 13-MAR-1995 (first entry)
XX
DE PtxA protein of Pasteurella haemolytica.
XX
KW Pasteurella haemolytica; shipping fever; pasteurellosis; cattle;
KW vaccine; treatment; therapy; prophylaxis; antigen; antibody;
KW immunoassay.
XX
OS Pasteurella haemolytica.
XX
PN US5336491-A.
XX
PD 09-AUG-1994.
XX

```
PF   15-JUN-1992;    92US-0899100.
XX
PR   28-NOV-1986;    86US-0935806.
PR   13-AUG-1987;    87US-0085430.
PR   18-JUN-1990;    90US-0540261.
PR   15-JUN-1992;    92US-0899100.
XX
PA   (TEXA ) UNIV TEXAS SYSTEM.
XX
PI   Berget P,  Engler M,  Highlander S,  Weinstock G;
XX
DR   WPI; 1994-255245/31.
DR   N-PSDB; AAQ70050.
XX
PT   New purified antigen from Pasteurella haemolytica - useful in
PT   vaccines against shipping fever of cattle, also for raising
PT   antibodies useful in diagnosis and passive immunisation
XX
PS   Disclosure; Figure 9; 35pp; English.
XX
CC   This protein is a 105 kD antigen which may be used in vaccines to
CC   protect cattle against shipping fever.  The antigen is also useful
CC   for raising antisera which can be used for passive immunisation (for
CC   treatment or temporary prophylaxis) and also to raise antibodies
CC   which can be used in immunoassay methods for the detection of
CC   Pasteurella haemolytica antigens in biological fluids.
CC   (Updated on 25-MAR-2003 to correct PF field.)
XX
SQ   Sequence   953 AA;

  Query Match             50.0%;  Score 2325;  DB 15;  Length 953;
  Best Local Similarity   49.9%;  Pred. No. 2.2e-136;
  Matches  464;  Conservative  175;  Mismatches  264;  Indels   26;  Gaps   13;

Qy          8 KSNIQAGLNSTKSGLKNLYLAIPKD--YDPQKGGTLNDFIKAADELGIARLAEEPNHTET 65

Db         28 QSLTQAG-SSLKTGAKKIILYIPQNYQYDTEQGNGLQDLVKAAEELGIEVQREERNNIAT 86

Qy         66 AKKSVDTVNQFLSLTQTGIAISATKLEKFLQKHSTNKLAKGLDSVENIDRKLGKASNVLS 125

Db         87 AQTSLGTIQTAIGLTERGIVLSAPQIDKLLQK---TKAGQALGSAESIVQNANKAKTVLS 143

Qy        126 TLSSFLGTALAGIELDSLIKKGDAAPDALAKASIDLINEIIGNLSQSTQTIEAFSSQLAK 185

Db        144 GIQSILGSVLAGMDLDEAL-QNNSNQHALAKAGLELTNSLIENIANSVKTLDEFGEQISQ 202

Qy        186 LGSTISQAKGFSNIGNKLQNL-NFSKTNLGLEIITGLLSGISAGFALADKNASTGKKVAA 244

Db        203 FGSKLQNIKGLGTLGDKLKNIGGLDKAGLGLDVISGLLSGATAALVLADKNASTAKKVGA 262

Qy        245 GFELSNQVIGNVTKAISSYVLAQRVAAGLSTTGAVAALITSSIMLAISPLAFMNAADKFN 304

Db        263 GFELANQVVGNITKAVSSYILAQRVAAGLSSTGPVAALIASTVSLAISPLAFAGIADKFN 322

Qy        305 HANALDEFAKQFRKFGYDGDHLLAEYQRGVGTIEASLTTISTALGAVSAGVSAAAVGSAV 364

Db        323 HAKSLESYAERFKKLGYDGDNLLAEYQRGTGTIDASVTAINTALAAIAGGVSAAAAGSVI 382

Qy        365 GTPIALLVAGVTGLISGILEASKQAMFESVANRLQGKILEWEKQNGGQNYFDKGYDSRYA 424

Db        383 ASPIALLVSGITGVISTILQYSKQAMFEHVANKIHNKIVEWEKNNHGKNYFENGYDARYL 442

Qy        425 AYLANNLKFLSELNKELEAERVIAITQQRWDNNIGELAGITKLGERIKSGKAYADAFEDG 484

Db        443 ANLQDNMKFLLNLNKELQAERVIAITQQQWDNNIGDLAGISRLGEKVLSGKAYVDAFEEG 502

Qy        485 KKVEAGSNITLDAKTGIIDISNSNGKKTQALHFTSPLLTAGTESRERLTNGKYSYINKLK 544

Db        503 KHIKADKLVQLDSANGIIDVSNSGKAKTQHILFRTPLLTPGTEHRERVQTGKYEYITKLN 562

Qy        545 FGRVKNWQVTDGEASSKLDFSKVIQRV-------AETEGTDEIGLIVNAKAGNDDIFVGQ 597

Db        563 INRVDSWKITDGAASSTFDLTNVVQRIGIELDNAGNVTKTKETKIIAKLGEGDDNVFVGS 622

Qy        598 GKMNIDGGDGHDRVFYSKDGGFGNITVDGTSATEAGSYTVNRKVARGDIYHEVVKRQETK 657

Db        623 GTTEIDGGEGYDRVHYSR-GNYGALTIDATKETEQGSYTVNRFVETGKALHEVTSTHTAL 681

Qy        658 VGKRTETIQYRDYELRKVGYGYQSTDNLKSVEEVIGSQFNDVFKGSKFNDIFHSGEGDDL 717

Db        682 VGNREEKIEYR-HSNNQHHAGYYTKDTLKAVEEIIGTSHNDIFKGSKFNDAFNGGDGVDT 740

Qy        718 LDGGAGDDRLFGGKGNDRLSGDEGDDLLDGGSGDDVLNGGAGNDVYIFRKGDGNDTLYDG 777

Db        741 IYGNDGNDRLFGGKGDDILDGGNGDDFIDGGKGNDLLHGGKGDDIFVHRKGDGNDIITDS 800

Qy        778 TGNDKLAFADANISDIMIERTKEGIIVKRNDHSGSINIPRWY----ITSNLQNYQSNKTD 833

Db        801 DGNDKLSFSDSNLKDLTFEKVKHNLVI-TNSKKEKVTIQNWFREADFAKEVPNYKATK-D 858

Qy        834 HKIEQLIGKDGSYITSDQIDKILQDKKDGTVITSQELKKLADENKSQKLSASDIASSLNK 893

Db        859 EKIEEIIGQNGERITSKQVDDLI--AKGNGKITQDELSKVVDNYELLKHS-KNVTNSLDK 915

Qy        894 LVGSMALFGTANSVSSNALQPITQPTQGI 922

Db        916 LISSVSAFTSSNDSRNVLVAPTSMLDQSL 944

RESULT 21
AAE04638
ID   AAE04638 standard; Protein; 953 AA.
XX
AC   AAE04638;
XX
DT   10-SEP-2001  (first entry)
XX
DE   Pasteurella haemolytica leukotoxin protein.
XX
KW   Leukotoxin; respiratory disease; infection; therapy; immunostimulant;
KW   antibacterial; vaccine; transgenic plant; animal feed.
XX
OS   Pasteurella haemolytica.
XX
FH   Key             Location/Qualifiers
FT   Domain          230..250
FT                   /label= Hydrophobic_transmembrane_domain_1
FT   Domain          280..320
FT                   /label= Hydrophobic_transmembrane_domain_2
FT   Domain          360..400
FT                   /label= Hydrophobic_transmembrane_domain_3
XX
PN   WO200144289-A2.
XX
PD   21-JUN-2001.
XX
PF   15-DEC-2000; 2000WO-CA01498.
XX
PR   17-DEC-1999;   99US-0172148.
XX
PA   (UYGU-) UNIV GUELPH OFFICE.
XX
PI   Lo RYC,  Shewen PE,  Lee RWH,  Hodgins D,  Strommer JN;
XX
DR   WPI; 2001-408470/43.
XX
PT   Modified leukotoxin polypeptide is useful in a vaccine to prevent or
PT   treat Mannheimia (Pasteurella) infection (particularly M. haemolytica
PT   infection), and disease associated with a leukotoxin, e.g., respiratory
PT   disease -
XX
PS   Claim 6; Fig 13; 70pp; English.
XX
CC   The present sequence is Pasteurella (Mannheimia) haemolytica full
CC   length leukotoxin protein. The present invention relates to modified
CC   leukotoxin DNA sequences, wherein the modification comprises the
```

```
CC   removal of DNA sequence encoding the amino acids within the hydrophobic
CC   transmembrane domain of a full length leukotoxin protein. Modified
CC   leukotoxin sequences are used in vaccines to treat or prevent diseases
CC   associated with leukotoxin, e.g., respiratory disease, and Mannheimia
CC   infection (particularly M. haemolytica infection). In addition, the
CC   vaccine is used to prepare a medicament. Furthermore, the plant
CC   transformed with modified leukotoxin sequences is fed to an animal
CC   such as a ruminant, to prevent or treat respiratory diseases.
XX
SQ   Sequence   953 AA;

  Query Match             50.0%;  Score 2325;  DB 22;  Length 953;
  Best Local Similarity   49.9%;  Pred. No. 2.2e-136;
  Matches  464;  Conservative  175;  Mismatches  264;  Indels   26;  Gaps   13;

Qy          8 KSNIQAGLNSTKSGLKNLYLAIPKD--YDPQKGGTLNDFIKAADELGIARLAEEPNHTET 65
              :|  ||| :| |:| | : | ||::  || ::|  | | :|||:||||     || |:  |
Db         28 QSLTQAG-SSLKTGAKKIILYIPQNYQYDTEQGNGLQDLVKAAEELGIEVQREERNNIAT 86

Qy         66 AKKSVDTVNQFLSLTQTGIAISATKLEKFLQKHSTNKLAKGLDSVENIDRKLGKASNVLS 125
              |: |: |:    : ||: || :|| :::| |||      |  : | | |:| :    ||  |||
Db         87 AQTSLGTIQTAIGLTERGIVLSAPQIDKLLQK---TKAGQALGSAESIVQNANKAKTVLS 143

Qy        126 TLSSFLGTALAGIELDSLIKKGDAAPDALAKASIDLINEIIGNLSQSTQTIEAFSSQLAK 185
               : | ||: |||::||  : : ::    |||||  ::| | :| |:: | :|:: |  |:::
Db        144 GIQSILGSVLAGMDLDEAL-QNNSNQHALAKAGLELTNSLIENIANSVKTLDEFGEQISQ 202

Qy        186 LGSTISQAKGFSNIGNKLQNL-NFSKTNLGLEIITGLLSGISAGFALADKNASTGKKVAA 244
               || :    ||     :|:||:|:      |  |||::|:||||| :     |||||||  ||| |
Db        203 FGSKLQNIKGLGTLGDKLKNIGGLDKAGLGLDVISGLLSGATAALVLADKNASTAKKVGA 262

Qy        245 GFELSNQVIGNVTKAISSYVLAQRVAAGLSTTGAVAALITSSIMLAISPLAFMNAADKFN 304
              |||:|||:||:|||:|||:||||||||||:|| ||||| |:: |||||||     |||||
Db        263 GFELANQVVGNITKAVSSYILAQRVAAGLSSTGPVAALIASTVSLAISPLAFAGIADKFN 322

Qy        305 HANALDEFAKQFRKFGYDGDHLLAEYQRGVGTIEASLTTISTALGAVSAGVSAAAVGSAV 364
              || :|: :|::|:| ||||:|||||||| |||:||:| |:||| |:: ||||| || :
Db        323 HAKSLESYAERFKKLGYDGDNLLAEYQRGTGTIDASVTAINTALAAIAGGVSAAAAGSVI 382

Qy        365 GTPIALLVAGVTGLISGILEASKQAMFESVANRLQGKILEWEKQNGGQNYFDKGYDSRYA 424
               :||||||:|:||:|| ||: |||||| |||::  ||:|||| | |:|||: |||:||
Db        383 ASPIALLVSGITGVISTILQYSKQAMFEHVANKIHNKIVEWEKNNHGKNYFENGYDARYL 442

Qy        425 AYLANNLKFLSELNKELEAERVIAITQQRWDNNIGELAGITKLGERIKSGKAYADAFEDG 484
              | | :|:|||  |||||:||||||||||:||||||:|||||:|||||:: ||||| ||||:|
Db        443 ANLQDNMKFLLNLNKELQAERVIAITQQQWDNNIGDLAGISRLGEKVLSGKAYVDAFEEG 502

Qy        485 KKVEAGSNITLDAKTGIIDISNSNGKKTQALHFTSPLLTAGTESRERLTNGKYSYINKLK 544
               | ::|    : ||:  ||||:|||      |||  : | :||||  ||| |||:  ||  || ||
Db        503 KHIKADKLVQLDSANGIIDVSNSGKAKTQHILFRTPLLTPGTEHRERVQTGKTEYITKLN 562

Qy        545 FGRVKNWQVTDGEASSKLDFSKVIQRV-------AETEGTDEIGLIVNAKAGNDDIFVGQ 597
                || :|::||| |||  | : |:||:           | |  :|      |:|::|||
Db        563 INRVDSWKITDGAASSTFDLTNVVQRIGIELDNAGNVTKTKETKIIAKLGEGDDNVFVGS 622

Qy        598 GKMNIDGGDGHDRVFYSKDGGFGNITVDGTSATEAGSYTVNRKVARGDIYHEVVKRQETK 657
              |     ||||:|:|||  ||:  |  :| :|:|  |   ||  |||||||  |    |     |||
Db        623 GTTEIDGGEGYDRVHYSR-GNYGALTIDATKETEQGSYTVNRFVETGKALHEVTSTHTAL 681

Qy        658 VGKRTETIQYRDYELRKVGYGYQSTDNLKSVEEVIGSQFNDVFKGSKFNDIFHSGEGDDL 717
              || | | |:|| :     :  || | ||:|||:||:    ||:|||||||||  |: |:| |
Db        682 VGNREEKIEYR-HSNNQHHAGYYTKDTLKAVEEIIGTSHNDIFKGSKFNDAFNGGDGVDT 740

Qy        718 LDGGAGDDRLFGGKGNDRLSGDEGDDLLDGGSGDDVLNGGAGNDVYIFRKGDGNDTLYDG 777
               :||  |:||||||||:| |  |||  |||  |:|:|||:|:::  ||||||:|  | : |
Db        741 IDGNDGNDRLFGGKGDDILDGGNGDDFIDGGKGNDLLHGGKGDDIFVHRKGDGNDIITDS 800

Qy        778 TGNDKLAFADANISDIMIERTKEGIIVKRNDHSGSINIPRWY----ITSNLQNYQSNKTD 833
                |||||:|:|:|: |:  |:  |   :::  |       :  |  |:     : ||:: |  |
Db        801 DGNDKLSFSDSNLKDLTFEKVKHNLVI-TNSKKEKVTIQNWFREADFAKEVPNYKATK-D 858

Qy        834 HKIEQLIGKDGSYITSDQIDKILQDKKDGTVITSQELKKLADENKSQKLSASDIASSLNK 893
               |||::||::|  |||  |:|  ::   |    ||  || |:  |  :  |  |   :: :||:|
Db        859 EKIEEIIGQNGERITSKQVDDLI--AKGNGKITQDELSKVVDNYELLKHS-KNVTNSLDK 915

Qy        894 LVGSMALFGTANSVSSNALQPITQPTQGI 922
              |: |:: | ::|    :  : | :    | :
Db        916 LISSVSAFTSSNDSRNVLVAPTSMLDQSL 944

RESULT 22
AAR34546
ID   AAR34546 standard; Protein; 943 AA.
XX
AC   AAR34546;
XX
DT   25-MAR-2003  (updated)
DT   23-AUG-1993  (first entry)
XX
DE   Somatostatin-leukotoxin gene fusion prod.
XX
KW   Vector; LKT 352; flanking; recombinant; antigen; somatostatin;
KW   gonadotropin releasing hormone; rotavirus viral protein 4;
KW   carrier protein; lactation; reproduction; SRIF.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Protein         1..929
FT                   /note= "recombinant leukotoxin protein"
FT   Peptide         930..943
FT                   /note= "SRIF"
XX
PN   WO9308290-A1.
XX
PD   29-APR-1993.
XX
PF   15-OCT-1992;   92WO-CA00449.
XX
PR   16-OCT-1991;   91US-0779171.
PR   14-OCT-1992;   92US-0960932.
XX
PA   (UYSA-) UNIV SASKATCHEWAN.
XX
PI   Hughes HPA,  Potter AA,  Redmond MJ;
XX
DR   WPI; 1993-152482/18.
DR   N-PSDB; AAQ41321.
XX
PT   Immunological carrier system with enhanced immunogenicity -
PT   comprises chimeric protein comprising leuco:toxin peptide or
PT   homologous protein fused to antigen esp. somatostatin or
PT   gonadotropin releasing hormone
XX
PS   Example 2; Fig 6; 95pp; English.
XX
CC   Oligonucleotides contg. sequences from bovine somatostatin (SRIF)
CC   gene were constructed on a Pharmacia Gene Assembler using standard
CC   phosporamidite chemistry.  The oligonucleotides were annealed and
CC   ligated into vector pAA352 (contg. the Pateurella heamolytica leuko-
CC   toxin gene) which had been digested with BamHI.  The ligated DNA was
CC   used to transform E. coli strain JM105.  Transformants contg. the
CC   oligonucleotide inserts were identified by restiction endonuclease
CC   mapping and the recombinant plasmid designated pAA496.  The
CC   chimeric protein produced from the plasmid works to bring about a
CC   larger immune response than the antigen alone, i.e. the leukotoxin
CC   works as a carrier protein.
CC   See also AAR34545-8.
CC   (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ   Sequence   943 AA;

  Query Match             50.0%;  Score 2323.5;  DB 14;  Length 943;
  Best Local Similarity   50.0%;  Pred. No. 2.7e-136;
  Matches  461;  Conservative  174;  Mismatches  262;  Indels   25;  Gaps   12;
```

```
Qy        15 LNSTKSGLKNLYLAIPKD--YDPQKGGTLNDFIKAADELGIARLAEEPNHTETAKKSVDT 72
             |:  |:| | : | ||::  || ::| | | :|||:||||     || |:  ||: |: |
Db         7 LSFPKTGAKKIILYIPQNYQYDTEQGNGLQDDVKAAEELGIEVQREERNNIATAQTSLGT 66

Qy        73 VNQFLSLTQTGIAISATKLEKFLQKHSTNKLAKGLDSVENIDRKLGKASNVLSTLSSFLG 132
             :   : ||: || :|| :::| |||    |  : | | |:| :   ||  ||| : | ||
Db        67 IQTAIGLTERGIVLSAPQIDKLLQK---TKAGQALGSAESIVQNANKAKTVLSGIQSILG 123

Qy       133 TALAGIELDSLIKKGDAAPDALAKASIDLINEIIGNLSQSTQTIEAFSSQLAKLGSTISQ 192
             : |||::||  : : ::   ||||| ::| | :| |:: | :|::  | |::: || :
Db       124 SVLAGMDLDEAL-QNNSNQHALAKAGLELTNSLIENIANSVKTLDEFGEQISQFGSKLQN 182

Qy       193 AKGFSNIGNKLQNL-NFSKTNLGLEIITGLLSGISAGFALADKNASTGKKVAAGFELSNQ 251
              ||    :|:||:|:     |  |||::|:||||| :|   ||||||| ||| |||||:||
Db       183 IKGLGTLGDKLKNIGGLDKAGLGLDVISGLLSGATAALVLADKNASTAKKVGAGFELANQ 242

Qy       252 VIGNVTKAISSYVLAQRVAAGLSTTGAVAALITSSIMLAISPLAFMNAADKFNHANALDE 311
             |:||:|||:|||:||||||||||:|| ||||| |:: |||||||    ||||||| :|:
Db       243 VVGNITKAVSSYILAQRVAAGLSSTGPVAALIASTVSLAISPLAFAGIADKFNHAKSLES 302

Qy       312 FAKQFRKFGYDGDHLLAEYQRGVGTIEASLTTISTALGAVSAGVSAAAVGSAVGTPIALL 371
             :|::|:| |||||:||||||||| |||:||:| |:||| |:: |||||| || : :|||||
Db       303 YAERFKKLGYDGDNLLAEYQRGTGTIDASVTAINTALAAIAGGVSAAAAGSVIASPIALL 362

Qy       372 VAGVTGLISGILEASKQAMFESVANRLQGKILEWEKQNGGQNYFDKGYDSRYAAYLANNL 431
             |:|:||:|| ||: ||||||| |||:    ||:|||| | |:|||: |||:||| | | |:
Db       363 VSGITGVISTILQYSKQAMFEHVANKIHNKIVEWEKNNHGKNYFENGYDARYLANLQDNM 422

Qy       432 KFLSELNKELEAERVIAITQQRWDNNIGELAGITKLGERIKSGKAYADAFEDGKKVEAGS 491
             |||   |||||:||||||||||:||||||:|||||::|||:: |||||  ||||:|| ::|
Db       423 KFLLNLNKELQAERVIAITQQQWDNNIGDLAGISRLGEKVLSGKAYVDAFEEGKHIKADK 482

Qy       492 NITLDAKTGIIDISNSNGKKTQALHFTSPLLTAGTESRERLTNGKYSYINKLKFGRVKNW 551
              : ||:  ||||:|||   ||| : | :|||| ||| |||:  ||| || ||    || :|
Db       483 LVQLDSANGIIDVSNSGKAKTQHILFRTPLLTPGTEHRERVQTGKYEYITKLNINRVDSW 542

Qy       552 QVTDGEASSKLDFSKVIQRV-------AETEGTDEIGLIVNAKAGNDDIFVGQGKMNIDG 604
             ::||| |||  | : |:||:              | |  :|      |:|::||| |    |||
Db       543 KITDGAASSTFDLTNVVQRIGIELDNAGNVTKTKETKIIAKLGEGDDNVFVGSGTTEIDG 602

Qy       605 GDGHDRVFYSKDGGFGNITVDGTSATEAGSYTVNRKVARGDIYHEVVKRQETKVGKRTET 664
             |:|:||| ||:  : :|:| |  || ||||||| | |||||||  | |:   ||  | |
Db       603 GEGYDRVHYSR-GNYGALTIDATKETEQGSYTVNRFVETGKALHEVTSTHTALVGNREEK 661

Qy       665 IQYRDYELRKVGYGYQSTDNLKSVEEVIGSQFNDVFKGSKFNDIFHSGEGDDLLDGGAGD 724
             |:||  :    :  || : | ||:|||:|||: ||:||||||||| |: |:| | :|| |:
Db       662 IEYR-HSNNQHHAGYYTKDTLKAVEEIIGTSHNDIFKGSKFNDAFNGGDGVDTIDGNDGN 720

Qy       725 DRLFGGKGNDRLSGDEGDDLLDGGSGDDVLNGGAGNDVYIFRKGDGNDTLYDGTGNDKLA 784
             |||||||:| | |  ||| :||| |:|:|:|| ||:|:::: ||||||| : |  |||||:
Db       721 DRLFGGKGDDILDGGNGDDFIDGGKGNDLLHGGKGDDIFVHRKGDGNDIITDSDGNDKLS 780

Qy       785 FADANISDIMIERTKEGIIVKRNDHSGSINIPRWY----ITSNLQNYQSNKTDHKIEQLI 840
             |:|:|: |: |: |  ::: |      : ||  |:          : ||:: | | |||::|
Db       781 FSDSNLKDLTFEKVKHNLVI-TNSKKEKVTIQNWFREADFAKEVPNYKATK-DEKIEEII 838

Qy       841 GKDGSYITSDQIDKILQDKKDGTVITSQELKKLADENKSQKLSASDIASSLNKLVGSMAL 900
             |::|  ||| |:|: ::   |     ||  || |: |   |   |  :: :|:|:|: |::
Db       839 GQNGERITSKQVDDLI--AKGNGKITQDELSKVVDNYELLKHS-KNVTNSLDKLISSVSA 895

Qy       901 FGTANSVSSNALQPITQPTQGI 922
             | ::|   :  : |  :  |  :
Db       896 FTSSNDSRNVLVAPTSMLDQSL 917
```

RESULT 23
AAR34547

```
ID   AAR34547 standard; Protein; 936 AA.
XX
AC   AAR34547;
XX
DT   25-MAR-2003  (updated)
DT   23-AUG-1993  (first entry)
XX
DE   GnRH-leukotoxin gene fusion prod.
XX
KW   Vector; LKT 352; flanking; recombinant; antigen; somatostatin;
KW   gonadotropin releasing hormone; rotavirus viral protein 4;
KW   carrier protein; lactation; reproduction; SRIF.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Protein         1..926
FT                   /note= "recombinant leukotoxin protein"
FT   Peptide         927..936
FT                   /note= "GnRH"
XX
PN   WO9308290-A1.
XX
PD   29-APR-1993.
XX
PF   15-OCT-1992;   92WO-CA00449.
XX
PR   16-OCT-1991;   91US-0779171.
PR   14-OCT-1992;   92US-0960932.
XX
PA   (UYSA-) UNIV SASKATCHEWAN.
XX
PI   Hughes HPA,  Potter AA,  Redmond MJ;
XX
DR   WPI; 1993-152482/18.
DR   N-PSDB; AAQ41322.
XX
PT   Immunological carrier system with enhanced immunogenicity -
PT   comprises chimeric protein comprising leuco:toxin peptide or
PT   homologous protein fused to antigen esp. somatostatin or
PT   gonadotropin releasing hormone
XX
PS   Example 2; Fig 8; 95pp; English.
XX
CC   Oligonucleotides contg. sequences from bovine gonadotrophin
CC   releasing hormone (GnRH) gene were constructed on a Pharmacia Gene
CC   Assembler using standard phosporamidite chemistry.  The oligo-
CC   nucleotides were annealed and ligated into vector pAA352 (contg.
CC   the Pateurella heamolytica leuko-toxin gene) which had been digested
CC   with BamHI.  The ligated DNA was used to transform E. coli strain
CC   MH3000.  Transformants contg. the oligonucleotide inserts were
CC   identified by restiction endonuclease mapping and the recombinant
CC   plasmid designated pAA502.  The chimeric protein produced from the
CC   plasmid works to bring about a larger immune response than the antigen
CC   alone, i.e. the leukotoxin works as a carrier protein.
CC   See also AAR34545-8.
CC   (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ   Sequence   936 AA;

  Query Match             49.9%;  Score 2318.5;  DB 14;  Length 936;
  Best Local Similarity   49.9%;  Pred. No. 5.6e-136;
  Matches  460;  Conservative  174;  Mismatches  263;  Indels   25;  Gaps   12;

Qy        15 LNSTKSGLKNLYLAIPKD--YDPQKGGTLNDFIKAADELGIARLAEEPNHTETAKKSVDT 72
             |:  |:| | : | ||::  || ::| | | :|||:||||     || |:  ||: |: |
Db         7 LSFPKTGAKKIILYIPQNYQYDTEQGNGLQDLVKAAEELGIEVQREERNNIATAQTSLGT 66

Qy        73 VNQFLSLTQTGIAISATKLEKFLQKHSTNKLAKGLDSVENIDRKLGKASNVLSTLSSFLG 132
             :   : ||: || :|| :::| |||    |  : | | |:| :   ||  ||| : | ||
Db        67 IQTAIGLTERGIVLSAPQIDKLLQK---TKAGQALGSAESIVQNANKAKTVLSGIQSILG 123

Qy       133 TALAGIELDSLIKKGDAAPDALAKASIDLINEIIGNLSQSTQTIEAFSSQLAKLGSTISQ 192
             : |||::||  : : ::   ||||| ::| | :| |:: | :|::  | |::: || :
Db       124 SVLAGMDLDEAL-QNNSNQHALAKAGLELTNSLIENIANSVKTLDEXCEQISQFGSKLQN 182

Qy       193 AKGFSNIGNKLQNL-NFSKTNLGLEIITGLLSGISAGFALADKNASTGKKVAAGFELSNQ 251
              ||    :|:||:|:     |  |||::|:||||| :|   ||||||| ||| |||||:||
Db       183 IKGLGTLGDKLKNIGGLDKAGLGLDVISGLLSGATAALVLADKNASTAKKVGAGFELANQ 242
```

```
Qy        252 VIGNVTKAISSYVLAQRVAAGLSTTGAVAALITSSIMLAISPLAFMNAADKFNHANALDE 311
Db        243 VVGNITKAVSSYILAQRVAAGLSSTGPVAALIASTVSLAISPLAFAGIADKFNHAKSLES 302

Qy        312 FAKQFRKFGYDGDHLLAEYQRGVGTIEASLTTISTALGAVSAGVSAAAVGSAVGTPIALL 371
Db        303 YAERFKKLGYDGDNLLAEYQRGTGTIDASVTAINTALAAIAGGVSAAAAGSVIASPIALL 362

Qy        372 VAGVTGLISGILEASKQAMFESVANRLQGKILEWEKQNGGQNYFDKGYDSRYAAYLANNL 431
Db        363 VSGITGVISTILQYSKQAMFEHVANKIHNKIVEWEKNNHGKNYFENGYDARYLANLQDNM 422

Qy        432 KFLSELNKELEAERVIAITQQRWDNNIGELAGITKLGERIKSGKAYADAFEDGKKVEAGS 491
Db        423 KFLLNLNKELQAERVIAITQQQWDNNIGDLAGISRLGEKVLSGKAYVDAFEEGKHIKADK 482

Qy        492 NITLDAKTGIIDISNSNGKKTQALHFTSPLLTAGTESRERLTNGKYSYINKLKFGRVKNW 551
Db        483 LVQLDSANGIIDVSNSGKAKTQHILFRTPLLTPGTEHRERVQTGKYEYITKLNINRVDSW 542

Qy        552 QVTDGEASSKLDFSKVIQRV-------AETEGTDEIGLIVNAKAGNDDIFVGQGKMNIDG 604
Db        543 KITDGAASSTFDLTNVVQRIGIELDNAGNVTKTKETKIIAKLGEGDDNVFVGSGTTEIDG 602

Qy        605 GDGHDRVFYSKDGGFGNITVDGTSATEAGSYTVNRKVARGDIYHEVVKRQETKVGKRTET 664
Db        603 GEGYDRVHYSR-GNYGALTIDATKETEQGSYTVNRFVETGKALHEVTSTHTALVGNREEK 661

Qy        665 IQYRDYELRKVGYGYQSTDNLKSVEEVIGSQFNDVFKGSKFNDIFHSGEGDDLLDGGAGD 724
Db        662 IEYR-HSNNQHHAGYYTKDTLKAVEEIIGTSHNDIFKGSKFNDAFNGGDGVDTIDGNDGN 720

Qy        725 DRLFGGKGNDRLSGDEGDDLLDGGSGDDVLNGGAGNDVYIFRKGDGNDTLYDGTGNDKLA 784
Db        721 DRLFGGKGDDILDGGNGDDFIDGGKGNDLLHGGKGDDIFVHRKGDGNDIITDSDGNDKLS 780

Qy        785 FADANISDIMIERTKEGIIVKRNDHSGSINIPRWY----ITSNLQNYQSNKTDHKIEQLI 840
Db        781 FSDSNLKDLTFEKVKHNLVI-TNSKKEKVTIQNWFREADFAKEVPNYKATK-DEKIEEII 838

Qy        841 GKDGSYITSDQIDKILQDKKDGTVITSQELKKLADENKSQKLSASDIASSLNKLVGSMAL 900
Db        839 GQNGERITSKQVDDLI--AKGNGKITQDELSKVVDNYELLKHS-KNVTNSLDKLISSVSA 895

Qy        901 FGTANSVSSNALQPITQPTQGI 922
Db        896 FTSSNDSRNVLVAPTSMLDQSL 917
```

RESULT 24
AAR10889
ID AAR10889 standard; Protein; 924 AA.
XX
AC AAR10889;
XX
DT 11-APR-1991 (first entry)
XX
DE Leukotoxin 352 encoded by plasmid pAA352.
XX
KW LKT; vaccine; antigen; respiratory disease; shipping fever pneumonia.
XX
OS Pasteurella haemolytica A1 strain B122.
XX
PN CA2014033-A.
XX
PD 07-OCT-1990.
XX
PF 06-APR-1990; 90CA-2014033.
XX
PR 07-APR-1989; 89US-0335018.
XX
PA (UYSA-) UNIV SASKATCHEWAN.
XX
PI Acres SD, Babiuk LA, Potter AA, Lawman MJP;
XX
DR WPI; 1991-000097/01.
DR N-PSDB; AAQ10272.
XX
PT Pasteurella haemolytica proteins and genes - used for producing
PT vaccines to protect animals esp. cattle from respiratory diseases
PT e.g. pneumonia.
XX
PS Claim 13; Fig 5; 87pp; English.
XX
CC Plasmid pAA352 is derived from pAA114, a clone isolated from a
CC genomic library of P. haemolytica. The protein, designated "new
CC leukotoxin" or "LKT 352" is 98% homologous to authentic leukotoxin.
CC LKT 352 and pref. antigenic fragments of it, can be used in
CC vaccines to protect cattle from respiratory diseases. They can also
CC be used to produce antibodies for immunoaffinity purificn. of
CC further proteins. [Fig. contg. sequence v. poor].
CC See also AAR10890, AAR10909, AAR10910 and AAQ10783.
XX
SQ Sequence 924 AA;

Query Match 49.9%; Score 2317.5; DB 12; Length 924;
Best Local Similarity 50.0%; Pred. No. 6.3e-136;
Matches 461; Conservative 174; Mismatches 260; Indels 27; Gaps 13;

```
Qy         15 LNSTKSGLKNLYLAIPKD--YDPQKGGTLNDFIKAADELGIARLAEEPNHTETAKKSVDT 72
Db          7 LSFPKTGAKKIILYIPQNYQYDTEQG--LQDLVKAAEELGIEVQREERNNIATAQTSLGT 64

Qy         73 VNQFLSLTQTGIAISATKLEKFLQKHSTNKLAKGLDSVENIDRKLGKASNVLSTLSSFLG 132
Db         65 IQTAIGLTERGIVLSAPQIDKLLQK---TKAGQALGSAESIVQNANKAKTVLSGIQSILG 121

Qy        133 TALAGIELDSLIKKGDAAPDALAKASIDLINEIIGNLSQSTQTIEAFSSQLAKLGSTISQ 192
Db        122 SVLAGMDLDEAL-QNNSNQHALAKAGLELTNSLIENIANSVKTLDEFGEQISQFGSKLQN 180

Qy        193 AKGFSNIGNKLQNL-NFSKTNLGLEIITGLLSGISAGFALADKNASTGKKVAAGFELSNQ 251
Db        181 IKGLGTLGDKLKNIGGLDKAGLGLDVISGLLSGATAALVLADKNASTAKKVGAGFELANQ 240

Qy        252 VIGNVTKAISSYVLAQRVAAGLSTTGAVAALITSSIMLAISPLAFMNAADKFNHANALDE 311
Db        241 VVGNITKAVSSYILAQRVAAGLSSTGPVAALIASTVSLAISPLAFAGIADKFNHAKSLES 300

Qy        312 FAKQFRKFGYDGDHLLAEYQRGVGTIEASLTTISTALGAVSAGVSAAAVGSAVGTPIALL 371
Db        301 YAERFKKLGYDGDNLLAEYQRGTGTIDASVTAINTALAAIAGGVSAAAAGSVIASPIALL 360

Qy        372 VAGVTGLISGILEASKQAMFESVANRLQGKILEWEKQNGGQNYFDKGYDSRYAAYLANNL 431
Db        361 VSGITGVISTILQYSKQAMFEHVANKIHNKIVEWEKNNHGKNYFENGYDARYLANLQDNM 420

Qy        432 KFLSELNKELEAERVIAITQQRWDNNIGELAGITKLGERIKSGKAYADAFEDGKKVEAGS 491
Db        421 KFLLNLNKELQAERVIAITQQQWDNNIGDLAGISRLGEKVLSGKAYVDAFEEGKHIKADK 480

Qy        492 NITLDAKTGIIDISNSNGKKTQALHFTSPLLTAGTESRERLTNGKYSYINKLKFGRVKNW 551
Db        481 LVQLDSANGIIDVSNSGKAKTQHILFRTPLLTPGTEHRERVQTGKYEYITKLNINRVDSW 540

Qy        552 QVTDGEASSKLDFSKVIQRV-------AETEGTDEIGLIVNAKAGNDDIFVGQGKMNIDG 604
Db        541 KITDGAASSTFDLTNVVQRIGIELDNAGNVTKTKETKIIAKLGEGDDNVFVGSGTTEIDG 600

Qy        605 GDGHDRVFYSKDGGFGNITVDGTSATEAGSYTVNRKVARGDIYHEVVKRQETKVGKRTET 664
Db        601 GEGYDRVHYSR-GNYGALTIDATKETEQGSYTVNRFVETGKALHEVTSTHTALVGNREEK 659

Qy        665 IQYRDYELRKVGYGYQSTDNLKSVEEVIGSQFNDVFKGSKFNDIFHSGEGDDLLDGGAGD 724
```

```
Db        660 IEYR-HSNNQHHAGYYTKDTLKAVEEIIGTSHNDIFKGSKFNDAFNGGDGVDTIDGNDGN 718

Qy        725 DRLFGGKGNDRLSGDEGDDLLDGGSGDDVLNGGAGNDVYIFRKGDGNDTLYDGTGNDKLA 784
Db        719 DRLFGGKGDDILDGGNGDDFIDGGKGNDLLHGGKGDDIFVHRKGDGNDIITDSDGNDKLS 778

Qy        785 FADANISDIMIERTKEGIIVKRNDHSGSINIPRWY----ITSNLQNYQSNKTDHKIEQLI 840
Db        779 FSDSNLKDLTFEKVKHNLVI-TNSKKEKVTIQNWFREADFAKEVPNYKATK-DEKIEEII 836

Qy        841 GKDGSYITSDQIDKILQDKKDGTVITSQELKKLADENKSQKLSASDIASSLNKLVGSMAL 900
Db        837 GQNGERITSKQVDDLI--AKGNGKITQDELSKVVDNYELLKHS-KNVTNSLDKLISSVSA 893

Qy        901 FGTANSVSSNALQPITQPTQGI 922
Db        894 FTSSNDSRNVLVAPTSMLDQSL 915
```

RESULT 25
AAW03945
ID   AAW03945 standard; Protein; 926 AA.
XX
AC   AAW03945;
XX
DT   20-NOV-1996 (first entry)
XX
DE   P. haemolytica truncated leukotoxin (LKT352).
XX
KW   Leukotoxin; LKT; gonadotropin-releasing hormone; GnRH;
KW   fusion protein; immunogen; vaccine; fertility control;
KW   contraceptive; sterilisation; plasmid pAA352.
XX
OS   Pasteurella haemolytica A1 strain B122.
XX
PN   WO9624675-A1.
XX
PD   15-AUG-1996.
XX
PF   24-JAN-1996;   96WO-CA00049.
XX
PR   10-FEB-1995;   95US-0387156.
XX
PA   (UYSA-) UNIV SASKATCHEWAN.
XX
PI   Manns JG, Potter AA;
XX
DR   WPI; 1996-384447/38.
DR   N-PSDB; AAT37179.
XX
PT   Gonadotropin-releasing hormone multimer fusion proteins - with
PT   leukotoxin polypeptide for increased immunogenicity, useful in
PT   antifertility vaccine prodn.
XX
PS   Example 1; Fig 3A-3I; 87pp; English.
XX
CC   A truncated leukotoxin (AAW03945), LKT 352, lacks the cytotoxic
CC   portion of the native protein from Pasteurella haemolytica. It is
CC   the product of plasmid pAA352 which carries a truncated lktA gene
CC   (AAT37179). A fusion protein (AAW03942) between LKT352 and a
CC   gonadotropin releasing hormone tetramer can be expressed in
CC   Escherichia coli. This is useful as a vaccine for fertility
CC   control, partic. immunological sterilisation of domestic or
CC   farm animals.
XX
SQ   Sequence   926 AA;

Query Match 49.9%; Score 2317.5; DB 17; Length 926;
Best Local Similarity 49.8%; Pred. No. 6.3e-136;
Matches 459; Conservative 175; Mismatches 263; Indels 25; Gaps 12;

```
Qy         15 LNSTKSGLKNLYLAIPKD--YDPQKGGTLNDFIKAADELGIARLAEEPNHTETAKKSVDT 72
Db          7 LSFPKTGAKKIILYIPQNYQYDTEQGNGLQDLVKAAEELGIEVQREERNNIATAQTSLGT 66

Qy         73 VNQFLSLTQTGIAISATKLEKFLQKHSTNKLAKGLDSVENIDRKLGKASNVLSTLSSFLG 132
Db         67 IQTAIGLTERGIVLSAPQIDKLLQK---TKAGQALGSAESIVQNANKAKTVLSGIQSILG 123

Qy        133 TALAGIELDSLIKKGDAAPDALAKASIDLINEIIGNLSQSTQTIEAFSSQLAKLGSTISQ 192
Db        124 SVLAGMDLDEAL-QNNSNQHALAKAGLELTNSLIENIANSVKTLDEFGEQISQFGSKLQN 182

Qy        193 AKGFSNIGNKLQNL-NFSKTNLGLEIITGLLSGISAGFALADKNASTGKKVAAGFELSNQ 251
Db        183 IKGLGTLGDKLKNIGGLDKAGLGLDVISGLLSGATAALVLADKNASTAKKVGAGFELANQ 242

Qy        252 VIGNVTKAISSYVLAQRVAAGLSTTGAVAALITSSIMLAISPLAFMNAADKFNHANALDE 311
Db        243 VVGNITKAVSSYILAQRVAAGLSSTGPVAALIASTVSLAISPLAFAGIADKFNHAKSLES 302

Qy        312 FAKQFRKFGYDGDHLLAEYQRGVGTIEASLTTISTALGAVSAGVSAAAVGSAVGTPIALL 371
Db        303 YAERFKKLGYDGDNLLAEYQRGTGTIDASVTAINTALAAIAGGVSAAAAGSVIASPIALL 362

Qy        372 VAGVTGLISGILEASKQAMFESVANRLQGKILEWEKQNGGQNYFDKGYDSRYAAYLANNL 431
Db        363 VSGITGVISTILQYSKQAMFEHVANKIHNKIVEWEKNNHGKNYFENGYDARYLANLQDNM 422

Qy        432 KFLSELNKELEAERVIAITQQRWDNNIGELAGITKLGERIKSGKAYADAFEDGKKVEAGS 491
Db        423 KFLLNLNKELQAERVIAITQQQWDNNIGDLAGISRLGEKVLSGKAYVDAFEEGKHIKADK 482

Qy        492 NITLDAKTGIIDISNSNGKKTQALHFTSPLLTAGTESRERLTNGKYSYINKLKFGRVKNW 551
Db        483 LVQLDSANGIIDVSNSGKAKTQHILFRTPLLTPGTEHRERVQTGKYEYITKLNINRVDSW 542

Qy        552 QVTDGEASSKLDFSKVIQRV-------AETEGTDEIGLIVNAKAGNDDIFVGQGKMNIDG 604
Db        543 KITDGAASSTFDLTNVVQRIGIELDNAGNVTKTKETKIIAKLGEGDDNVFVGSGTTEIDG 602

Qy        605 GDGHDRVFYSKDGGFGNITVDGTSATEAGSYTVNRKVARGDIYHEVVKRQETKVGKRTET 664
Db        603 GEGYDRVHYSR-GNYGALTIDATKETEQGSYTVNRFVETGKALHEVTSTHTALVGNREEK 661

Qy        665 IQYRDYELRKVGYGYQSTDNLKSVEEVIGSQFNDVFKGSKFNDIFHSGEGDDLLDGGAGD 724
Db        662 IEYR-HSNNQHHAGYYTKDTLKAVEQIIGTSHNDIFKGSKFNDAFNGGDAVDTIDGNDGN 720

Qy        725 DRLFGGKGNDRLSGDEGDDLLDGGSGDDVLNGGAGNDVYIFRKGDGNDTLYDGTGNDKLA 784
Db        721 DRLFGGKGDDILDGGNGDDFIDGGKGNDLLHGGKGDDIFVHRKGDGNDIITDSDGNDKLS 780

Qy        785 FADANISDIMIERTKEGIIVKRNDHSGSINIPRWY----ITSNLQNYQSNKTDHKIEQLI 840
Db        781 FSDSNLKDLTFEKVKHNLVI-TNSKKEKVTIQNWFREADFAKEVPNYKATK-DEKIEEII 838

Qy        841 GKDGSYITSDQIDKILQDKKDGTVITSQELKKLADENKSQKLSASDIASSLNKLVGSMAL 900
Db        839 GQNGERITSKQVDDLI--AKGNGKITQDELSKVVDNYELLKHS-KNVTNSLDKLISSVSA 895

Qy        901 FGTANSVSSNALQPITQPTQGI 922
Db        896 FTSSNDSRNVLVAPTSMLDQSL 917
```

RESULT 26
AAR15159
ID   AAR15159 standard; Protein; 953 AA.
XX
AC   AAR15159;
XX
DT   10-JAN-1992 (first entry)
XX
DE   Leukotoxin from P. haemolytica.
XX

```
KW   Leukotoxin; probe; pasteurellosis; cytotoxicity; leukocytes.
XX
OS   Pasteurella haemolytica A1 (ATCC 43270).
XX
PN   US5055400-A.
XX
PD   08-OCT-1991.
XX
PF   26-NOV-1986;   86US-0935493.
XX
PR   26-NOV-1986;   86US-0935493.
XX
PA   (UYGU-) UNIV OF GUELPH.
XX
PI   Lo RYC,  Shewen PE,  Strathdee CA;
XX
DR   WPI; 1991-317648/43.
DR   N-PSDB; AAQ14000.
XX
PT   DNA encoding Pasteurella haemolytica-A1 leukotoxin - used to
PT   produce recombinant protein for prepn. of vaccine and antibodies
PT   for protection of cattle from pneumonic pasteurellosis
XX
PS   Claim 6; Page 15; 20pp; English.
XX
CC   ORF2 (AAR14223) codes for a 19.820 D protein of 166 amino acids;
CC   ORF1 codes for leukotoxin, a 101.883 D protein of 952 amino acids,
CC   which is produced on expression of pLKT5 in a host organism.
CC   The leukotoxin has cytotoxic activity specific against leukocytes.
CC   The protein is used to prepare vaccines or to raise antibodies
CC   against P. haemolytica to protect cattle from pneumonic pasteurellosis.
XX
SQ   Sequence   953 AA;

  Query Match             49.8%;  Score 2313;  DB 12;  Length 953;
  Best Local Similarity   49.6%;  Pred. No. 1.3e-135;
  Matches  461;  Conservative  176;  Mismatches  266;  Indels   26;  Gaps   13;

Qy          8 KSNIQAGLNSTKSGLKNLYLAIPKD--YDPQKGGTLNDFIKAADELGIARLAEEPNHTET 65
              :|  |||  :|  |:|  |  :  |  ||::  ||  ::|  |  |  :|||:||||    ||  |:  |
Db         28 QSLTQAG-SSLKTGAKKIILYIPQNYQYDTEQGNGLQDLVKAAEELGIEVQREERNNIAT 86

Qy         66 AKKSVDTVNQFLSLTQTGIAISATKLEKFLQKHSTNKLAKGLDSVENIDRKLGKASNVLS 125
              |:  |:  |:    :  ||:  ||  :||  :::|  |||    |  :  |  |  |:|  :    ||  |||
Db         87 AQTSLGTIQTAIGLTERGIVLSAPQIDKLLQK---TKAGQALGSAESIVQNANKAKTVLS 143

Qy        126 TLSSFLGTALAGIELDSLIKKGDAAPDALAKASIDLINEIIGNLSQSTQTIEAFSSQLAK 185
               :  |  ||:  |||::||    :  :  ::      |||||  ::|  |  :|  |::  |  :|::  |    |:::
Db        144 GIQSILGSVLAGMDLDEAL-QNNSNQHALAKAGLELTNSLIENIANSVKTLDEFGEQISQ 202

Qy        186 LGSTISQAKGFSNIGNKLQNL-NFSKTNLGLEIITGLLSGISAGFALADKNASTGKKVAA 244
               ||  :    ||      :|:||:|:      |   |||::|:||||  :|   ||||||||  |||  |
Db        203 FGSKLQNIKGLGTLGDKLKNIGGLDKAGLGLDVISGLLSGATAALVLADKNASTAKKVGA 262

Qy        245 GFELSNQVIGNVTKAISSYVLAQRVAAGLSTTGAVAALITSSIMLAISPLAFMNAADKFN 304
              ||||:||::||:|||:|||:||||||||||:||  |||||  |::  ||||||||      |||||
Db        263 GFELANQLVGNITKAVSSYILAQRVAAGLSSTGPVAALIASTVSLAISPLAFAGIADKFN 322

Qy        305 HANALDEFAKQFRKFGYDGDHLLAEYQRGVGTIEASLTTISTALGAVSAGVSAAAVGSAV 364
              ||  :|:  :|::|:|  |||||:|||||||  |||:||:|  |:|||  |::  ||||||  ||  :
Db        323 HAKSLESYAERFKKLGYDGDNLLAEYQRGTGTIDASVTAINTALAAIAGGVSAAAAGSVI 382

Qy        365 GTPIALLVAGVTGLISGILEASKQAMFESVANRLQGKILEWEKQNGGQNYFDKGYDSRYA 424
               :|||||:|:||:||  ||:  |||||||  |||::  |:||||  |  |:|||:  |||:|
Db        383 ASPIALLVSGITGVISTILQYSKQAMFEHVANKIHNKIVEWEKNNHGKNYFENGYDARYL 442

Qy        425 AYLANNLKFLSELNKELEAERVIAITQQRWDNNIGELAGITKLGERIKSGKAYADAFEDG 484
              |  |  :|:|||    |||||:||  |||||||:|||||||:|||||:||:  |||||        ||:|
Db        443 ANLQDNMKFLLNLNKELQAEAVIAITQQQWDNNIGDLAGISRLGEKVLSGKAYVHPFEEG 502

Qy        485 KKVEAGSNITLDAKTGIIDISNSNGKKTQALHFTSPLLTAGTESRERLTNGKYSYINKLK 544
              |  ::|    :  ||:  ||||:|||  ||||||:|  |||  :  :||||  |||  |||:  |||  ||  ||
Db        503 KHIKADKLVQLDSANGIIDVSNSGKAKTQHILFRTPLLTPGTEHRERVQTGKYEYITKLN 562

Qy        545 FGRVKNWQVTDGEASSKLDFSKVIQRV-------AETEGTDEIGLIVNAKAGNDDIFVGQ 597
              ||  :|::|||  |||    |  :  |:||:                  |  |    :|          |:|::|||
Db        563 INRVDSWKITDGAASSTFDLTNVVQRIGIELDNAGNVTKTKETKIIAKLGEGDDNVFVGS 622

Qy        598 GKMNIDGGDGHDRVFYSKDGGFGNITVDGTSATEAGSYTVNRKVARGDIYHEVVKRQETK 657
              |      ||||:|:|||  ||:  |  :|  :|:|  |    ||  ||||||  |      |      |||
Db        623 GTTEIDGGEGYDRVHYSR-GNYGALTIDATKETEQGSYTVNRFVETGKALHEVTSTHTAL 681

Qy        658 VGKRTETIQYRDYELRKVGYGYQSTDNLKSVEEVIGSQFNDVFKGSKFNDIFHSGEGDDL 717
              ||  |  |  |:||  :      :      ||  :  |  ||:|||:||:    ||:||||||||  |:  |:|  |
Db        682 VGNREEKIEYR-HSNNQHHAGYYTKDTLKAVEEIIGTSHNDIFKGSKFNDAFNGGDGVDT 740

Qy        718 LDGGAGDDRLFGGKGNDRLSGDEGDDLLDGGSGDDVLNGGAGNDVYIFRKGDGNDTLYDG 777
              :||    |:||||||||:|  |  |      |||  :|||  |:|:|:||  |:|:::  ||||||  :  |
Db        741 IDGNDGNDRLFGGKGDDILDGGNGDDFIDGGKGNDLLHGGKGDDIFVHRKGDGNDIITDS 800

Qy        778 TGNDKLAFADANISDIMIERTKEGIIVKRNDHSGSINIPRWY----ITSNLQNYQSNKTD 833
              |||||:|:|:|:  |:    |:  |      |:      :::    |  ·      :      |      :  :  ||::  |  |
Db        801 DGNDKLSFSDSNLKDLTFEKVKHNLVI-TNSKKEKVTIQNWFREADFAKEVPNYKATK-D 858

Qy        834 HKIEQLIGKDGSYITSDQIDKILQDKKDGTVITSQELKKLADENKSQKLSASDIASSLNK 893
              |||::||::|    |||  |:|:  ::          ||      ||  ||  |:    |        :      |    |    ::  :||:|
Db        859 EKIEEIIGQNGERITSKQVDDLI--AKGNGKITQDELSKVVDNYELLKHS-KNVTNSLDK 915

Qy        894 LVGSMALFGTANSVSSNALQPITQPTQGI 922
              |:  |::  |  ::|      :    :  |  :      |  :
Db        916 LISSVSAFTSSNDSRNVLVAPTSMLDQSL 944

RESULT 27
AAR34548
ID   AAR34548 standard; Protein; 951 AA.
XX
AC   AAR34548;
XX
DT   25-MAR-2003  (updated)
DT   23-AUG-1993  (first entry)
XX
DE   Rotavirus VP4-leukotoxin gene fusion prod.
XX
KW   Vector; LKT 352; flanking; recombinant; antigen; somatostatin;
KW   gonadotropin releasing hormone; rotavirus viral protein 4;
KW   carrier protein; lactation; reproduction; SRIF.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Protein         1..926
FT                   /note= "recombinant leukotoxin protein"
FT   Peptide         927..951
FT                   /note= "Rotavirus VP-4"
XX
PN   WO9308290-A1.
XX
PD   29-APR-1993.
XX
PF   15-OCT-1992;   92WO-CA00449.
XX
PR   16-OCT-1991;   91US-0779171.
PR   14-OCT-1992;   92US-0960932.
XX
PA   (UYSA-) UNIV SASKATCHEWAN.
XX
PI   Hughes HPA,  Potter AA,  Redmond MJ;
XX
DR   WPI; 1993-152482/18.
DR.  N-PSDB; AAQ41322.
XX
PT   Immunological carrier system with enhanced immunogenicity -
PT   comprises chimeric protein comprising leuco:toxin peptide or
PT   homologous protein fused to antigen esp. somatostatin or
PT   gonadotropin releasing hormone
```

```
XX
PS   Example 2; Fig 8; 95pp; English.
XX
CC   Oligonucleotides contg. sequences from bovine Rotavirus viral
CC   protein 4 (VP-4) gene were constructed on a Pharmacia Gene
CC   Assembler using standard phosporamidite chemistry.  The oligo-
CC   nucleotides were annealed and ligated into vector pAA352 (contg.
CC   the Pateurella heamolytica leuko-toxin gene) which had been digested
CC   with BamHI.  The ligated DNA was used to transform E. coli strain
CC   MH3000.  Transformants contg. the oligonucleotide inserts were
CC   identified by restiction endonuclease mapping and the recombinant
CC   plasmid designated pAA501.  The chimeric protein produced from the
CC   plasmid works to bring about a larger immune response than the antigen
CC   alone, i.e. the leukotoxin works as a carrier protein.
CC   See also AAR34545-7.
CC   (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ   Sequence   951 AA;

  Query Match            49.8%;  Score 2312.5;  DB 14;  Length 951;
  Best Local Similarity  49.8%;  Pred. No. 1.3e-135;
  Matches  459;  Conservative  174;  Mismatches  264;  Indels   25;  Gaps   12;

Qy         15 LNSTKSGLKNLYLAIPKD--YDPQKGGTLNDFIKAADELGIARLAEEPNHTETAKKSVDT 72
              |:  |:|  | :  | ||::  || ::|  |  | :|||:||||     || |:  ||: |: |
Db          7 LSFPKTGAKKIILYIPQNYQYDTEQGNGLQDLVKAAEELGIEVQREERNNIATAQTSLGT 66

Qy         73 VNQFLSLTQTGIAISATKLEKFLQKHSTNKLAKGLDSVENIDRKLGKASNVLSTLSSFLG 132
              :   :  ||: || :|| :::| |||      |  :  |  ||  |:|  :    ||  |||  |  ||
Db         67 IQTAIGLTERGIVLSAPQIDKLLQK---TKAGQALGSAESIVQNANKAKTVLSGIQSILG 123

Qy        133 TALAGIELDSLIKKGDAAPDALAKASIDLINEIIGNLSQSTQTIEAFSSQLAKLGSTISQ 192
              : |||::||   : : ::     ||||  ::|  |  :|  |:: | :|::    |::: || :
Db        124 SVLAGMDLDEAL-QNNSNQHALAKAGLELTNSLIENIANSVKTLDEXCEQISQFGSKLQN 182

Qy        193 AKGFSNIGNKLQNL-NFSKTNLGLEIITGLLSGISAGFALADKNASTGKKVAAGFELSNQ 251
               ||     :|:||:|:         |   |||::|:||||  :|     |||||||| ||| |||||:||
Db        183 IKGLGTLGDKLKNIGGLDKAGLGLDVISGLLSGATAALVLADKNASTAKKVGAGFELANQ 242

Qy        252 VIGNVTKAISSYVLAQRVAAGLSTTGAVAALITSSIMLAISPLAFMNAADKFNHANALDE 311
              |:||:|||:|||:||||||||||:||  ||||| |:: |||||||    ||||||| :|:
Db        243 VVGNITKAVSSYILAQRVAAGLSSTGPVAALIASTVSLAISPLAFAGIADKFNHAKSLES 302

Qy        312 FAKQFRKFGYDGDHLLAEYQRGVGTIEASLTTISTALGAVSAGVSAAAVGSAVGTPIALL 371
              :|::|:|  |||||:||||||||  |||:||:|  |:|||  |::  |||||  || :  :|||||
Db        303 YAERFKKLGYDGDNLLAEYQRGTGTIDASVTAINTALAAIAGGVSAAAAGSVIASPIALL 362

Qy        372 VAGVTGLISGILEASKQAMFESVANRLQGKILEWEKQNGGQNYFDKGYDSRYAAYLANNL 431
              |:|:||:||  ||:|| |||||||| |||:    ||:|||| |  |:|||:  |||:|| |  |  :|:
Db        363 VSGITGVISTILQYSKQAMFEHVANKIHNKIVEWEKNNHGKNYFENGYDARYLANLQDNM 422

Qy        432 KFLSELNKELEAERVIAITQQRWDNNIGELAGITKLGERIKSGKAYADAFEDGKKVEAGS 491
              |||  |||||:||||||||||:|||  ||:||||::|||::  |||||  ||||:|| ::|
Db        423 KFLLNLNKELQAERVIAITQQQWDNKIGDLAGISRLGEKVLSGKAYVDAFEEGKHIKADK 482

Qy        492 NITLDAKTGIIDISNSNGKKTQALHFTSPLLTAGTESRERLTNGKYSYINKLKFGRVKNW 551
               : ||:  |||:|||   |||  :  | :||||  ||| |||:  |||  || ||    || :|
Db        483 LVQLDSANGIIDVSNSGKAKTQHILFRTPLLTPGTEHRERVQTGKYEYITKLNINRVDSW 542

Qy        552 QVTDGEASSKLDFSKVIQRV-------AETEGTDEIGLIVNAKAGNDDIFVGQGKMNIDG 604
              ::|||  |||  |  |:||:                 |  |   :|      |:|::|||  |    |||
Db        543 KITDGAASSTFDLTNVVQRIGIELDNAGNVTKTKETKIIAKLGEGDDNVFVGSGTTEIDG 602

Qy        605 GDGHDRVFYSKDGGFGNITVDGTSATEAGSYTVNRKVARGDIYHEVVKRQETKVGKRTET 664
              |:|:||| ||:  |  :|  :|:|  |   ||  |||||||  |             |||        ||  |  |
Db        603 GEGYDRVHYSR-GNYGALTIDATKETEQGSYTVNRFVETGKALHEVTSTHTALVGNREEK 661

Qy        665 IQYRDYELRKVGYGYQSTDNLKSVEEVIGSQFNDVFKGSKFNDIFHSGEGDDLLDGGAGD 724
              |:||  :     :      ||  :  |  ||:|||:||:  ||:||||||||  |: |:|  |  :||  |:
Db        662 IEYR-HSNNQHHAGYYTKDTLKAVEEIIGTSHNDIFKGSKFNDAFNGGDGVDTIDGNDGN 720

Qy        725 DRLFGGKGNDRLSGDEGDDLLDGGSGDDVLNGGAGNDVYIFRKGDGNDTLYDGTGNDKLA 784
              ||||||||:|  |  |  |||  :|||  |:|:|:|:||  |:|:::  ||||||  .:  |  |||||:
Db        721 DRLFGGKGDDILDGGNGDDFIDGGKGNDLLHGGKGDDIFVHRKGDGNDIITDSDGNDKLS 780

Qy        785 FADANISDIMIERTKEGIIVKRNDHSGSINIPRWY----ITSNLQNYQSNKTDHKIEQLI 840
              |:|:|:  |:   |:  |  :::  |        :  |  |:          :  ||::  |  |  |||::|
Db        781 FSDSNLKDLTFEKVKHNLVI-TNSKKEKVTIQNWFREADFAKEVPNYKATK-DEKIEEII 838

Qy        841 GKDGSYITSDQIDKILQDKKDGTVITSQELKKLADENKSQKLSASDIASSLNKLVGSMAL 900
              |::|  |||  |:|  ::     |       ||   ||  |:  |    :  |  |     ::  :||:||:  |::
Db        839 GQNGERITSKQVDDLI--AKGNGKITQDELSKVVDNYELLKHS-KNVTNSLDKLISSVSA 895

Qy        901 FGTANSVSSNALQPITQPTQGI 922
              |  ::|      :   :  |  :      |  :
Db        896 FTSSNDSRNVLVAPTSMLDQSL 917


RESULT 28
AAW07637
ID   AAW07637 standard; Protein; 934 AA.
XX
AC   AAW07637;
XX
DT   25-MAR-2003  (updated)
DT   05-FEB-1997  (first entry)
XX
DE   P. suis leukotoxin gene pslktA, protein product.
XX
KW   Leukotoxin; pslktC; pslktA; pslktB; calcium dependent; RTX;
KW   cytotoxic protein; repeat of toxin; BL-3 cell; pig; lymphocyte;
KW   immunogen; vaccine; immunisation; diarrhoea; reagent;
KW   diagnosis; assay; enteritis.
XX
OS   Pasteurella suis.
XX
FH   Key             Location/Qualifiers
FT   Domain          229..245
FT                   /label= transmembrane_domain
FT   Domain          295..311
FT                   /label= transmembrane_domain
FT   Domain          377..393
FT                   /label= transmembrane_domain
FT   Region          710..718
FT                   /note= "glycine rich repeat sequence"
FT   Region          719..727
FT                   /note= "glycine rich repeat sequence"
FT   Region          728..736
FT                   /note= "glycine rich repeat sequence"
FT   Region          737..745
FT                   /note= "glycine rich repeat sequence"
FT   Region          746..754
FT                   /note= "glycine rich repeat sequence"
FT   Region          756..763
FT                   /note= "glycine rich repeat sequence"
FT   Region          764..772
FT                   /note= "glycine rich repeat sequence"
FT   Region          773..781
FT                   /note= "glycine rich repeat sequence"
XX
PN   US5559008-A.
XX
PD   24-SEP-1996.
XX
PF   22-MAR-1994;   94US-0215805.
XX
PR   05-NOV-1992;   92US-0972157.
PR   22-MAR-1994;   94US-0215805.
XX
PA   (CORR ) CORNELL RES FOUND INC.
XX
PI   Chang Y;
XX
DR   WPI; 1996-442408/44.
DR   N-PSDB; AAT45417.
XX
```

```
PT   Leukotoxin genes from Pasteurella suis - also vectors and cells for
PT   expressing gene prods., for use in vaccines against porcine
PT   enteritis
XX
PS   Claim 34; Columns 59-64; 47pp; English.
XX
CC   The present sequence is protein prod. of the Pasteurella suis
CC   leukotoxin gene pslktA, which together with the pslktC, and pslktB
CC   gene prods., comprises a high mol. wt. calcium dependent cytotoxic
CC   protein of the repeat of toxin family (RTX), capable of killing
CC   both BL-3 cells and pig lymphocytes. The leukotoxin proteins, and
CC   the pslktC, A and B genes, or fragments, can be used as immunogens
CC   in vaccines for the immunisation of pigs against diarrhoea, or as
CC   reagents in diagnostic assays. The leukotoxin genes were isolated
CC   by screening a P. suis genomic library, in the phage vector
CC   lambda-dash, with a DNA probe derived from pYFC19, a plasmid
CC   carrying the IKtCA locus (Chang et al., Infect. Immun., vol. 55,
CC   pp2348-2354, 1987).
CC   (Updated on 25-MAR-2003 to correct PF field.)
XX
SQ   Sequence   934 AA;

  Query Match             49.4%;  Score 2297.5;  DB 17;  Length 934;
  Best Local Similarity   49.7%;  Pred. No. 1.1e-134;
  Matches  469;  Conservative  172;  Mismatches  254;  Indels   49;  Gaps   17;

Qy          1 MSNINV-IKSNIQAGLNST-----------KSGLKNLYLAIPKD--YDPQKGGTLNDFIK 46
Db          4 LANISTNLKNSLQSGLHKTGQSLNQAGQSLKAGAKKLILYIPKDYEYDSGRGNGLQDLVK 63

Qy         47 AADELGIARLAEEPNHTETAKKSVDTVNQFLSLTQTGIAISATKLEKFLQKHSTNKLAKG 106
Db         64 AAEDLGIEVQREERNGIATAQNSLSTIQNILGFSERGVVLSAPQLDKLLQKY---KISKA 120

Qy        107 LDSVENIDRKLGKASNVLSTLSSFLGTALAGIELDSLIKKGDAAPDALAKASIDLINEII 166
Db        121 PGSSENVAKNLGNAQTLLSGIQSILGSVMAGMDLDEILKNKGSELD-LAKAGLELTNSLI 179

Qy        167 GNLSQSTQTIEAFSSQLAKLGSTISQAKGFSNIGNKLQNLN-FSKTNLGLEIITGLLSGI 225
Db        180 ENIANSVQTLDTFSEQISQLGTKLQNVKGLGTLGDKLKNFSGFSKAGLGLEVISGLLSGA 239

Qy        226 SAGFALADKNASTGKKVAAGFELSNQVIGNVTKAISSYVLAQRVAAGLSTTGAVAALITS 285
Db        240 TAALVLADKNASTDRKVGAGFELANQVVGNITKAVSSYILAQRVAAGLSNTGPVSALIAS 299

Qy        286 SIMLAISPLAFMNAADKFNHANALDEFAKQFRKFGYDGDHLLAEYQRGVGTIEASLTTIS 345
Db        300 TVALAISPLAFAGIADKFNNAKALESYAERFKKLGYEGDSLLAEYQRGTGTIDASVTAVN 359

Qy        346 TALGAVSAGVSAAAVGSAVGTPIALLVAGVTGLISGILEASKQAMFESVANRLQGKILEW 405
Db        360 TALAAISGGVSAAAAGSLVGAPIALLVSGITGIISTILQYSKQAMFEHVANKIHDKIVDW 419

Qy        406 EKQNGGQNYFDKGYDSRYAAYLANNLKFLSELNKELEAERVIAITQQRWDNNIGELAGIT 465
Db        420 EKKHNGKNYFENGYDSRYLADLQDNMRQLQNLNKELQAERVIRITQQQWDNNIGNLAGIS 479

Qy        466 KLGERIKSGKAYADAFEDGKKVEAGSNITLDAKTGIIDISNSNGKKTQALHFTSPLLTAG 525
Db        480 RLGEKVMSGKAYADAFEEGKLIKADTFVQLDSATGVINTSKSDNVKTQHILFRTPLLTPG 539

Qy        526 TESRERLTNGKYSYINKLKFGRVKNWQVTDGEASSKLDFSKVIQRVA-------ETEGTD 578
Db        540 VENRERIQTGKYEYITKLNINRVDSWKITDGATNSTFDLTNVVQRIGIELDHADNVTKTK 599

Qy        579 EIGLIVNAKAGNDDIFVGQGKMNIDGGDGHDRVFYSKDGGFGNITVDGTSATEAGSYTVN 638
Db        600 ETKIIANLGDGNDDVFIGSGTTEVDGGNGLDRVHYSR-GDYGALTIDATNESVQGSYTVK 658

Qy        639 RKVARGDIYHEVVKRQETKVGKRTETIQYRDYELRKVGYGYQSTDNLKSVEEVIGSQFND 698
Db        659 RFVETGKALHEVTATQSVLVGSREEKIEYR-HSNNTQHAGYYTTDTLKSVEEIIGTSRND 717

Qy        699 VFKGSKFNDIFHSGEGDDLLDGGAGDDRLFGGKGNDRLSGDEGDDLLDGGSGDDVLNGGA 758
Db        718 IFKGSKFDDAFHGGDGVDNIDGNAGNDRLFGGKGFDIIDGGDGDDFIDGGQGDDILHGGK 777

Qy        759 GNDVYIFRKGDGNDTLYDGTGNDKLAFADANISDIMIERTKEGIIVKRNDHSGSINIPRW 818
Db        778 GNDILCTVKG-GNDSISDSGGNDRLSFADSNLKDLTFEKVNHHLMI-TNVKKEKVTIQNW 835

Qy        819 Y----ITSNLQNYQSNKTDHKIEQLIGKDGSYITSDQIDKILQDKKDGTVITSQELKKLA 874
Db        836 FREADYAKTVHNYQAT-ADEKIEEIIGRQGERITSKQIDELIEKGKG--KIDQSELERIA 892

Qy        875 DENKSQKLSASDIAS-SLNKLVGSMALFGTANSVSSNALQPITQ 917
Db        893 E--SSALLKESKFASNSLNKLVSSAAHL---------PLQTITE 925


RESULT 29
AAR12561
ID   AAR12561 standard; Protein; 956 AA.
XX
AC   AAR12561;
XX
DT   25-MAR-2003  (updated)
DT   26-JUL-1991  (first entry)
XX
DE   APPA haemolysin antigen encoded by appA.
XX
KW   Vaccine; porcine pleuroneumonia; haemolysin; cytotoxin.
XX
OS   Actinobacillus pleuroneumoniae serotype 5.
XX
PN   WO9106653-A.
XX
PD   16-MAY-1991.
XX
PF   31-OCT-1990;   90WO-US06350.
XX
PR   31-OCT-1989;   89US-0429273.
XX
PA   (TEXA ) UNIV TEXAS A & M SYSTEM.
XX
PI   Struck DK,  Young RF,  Chang YF;
XX
DR   WPI; 1991-164202/22.
DR   N-PSDB; AAQ11873.
XX
PT   DNA encoding Actinobacillus pleuropneumoniae haemolysin - used
PT   for producing haemolysin antigen for vaccinating pigs against
PT   porcine pleuro-pneumonia.
XX
PS   Claim 8; Fig 1; 28pp; English.
XX
CC   The sequence is encoded by the appA gene carried on a 3.8 kb
CC   section of clone yfc5 which was isolated by immunological screen-
CC   ing. It forms the appCA antigen with the appC protein encoded on
CC   the same section of DNA.  The appA antigen is not in itself
CC   haemolytic but does elicit an immune response and can be used in
CC   vaccines.
CC   See also AAR12127.
CC   (Updated on 25-MAR-2003 to correct PA field.)
XX
SQ   Sequence   956 AA;

  Query Match             48.1%;  Score 2236.5;  DB 12;  Length 956;
  Best Local Similarity   47.7%;  Pred. No. 7.4e-131;
  Matches  453;  Conservative  188;  Mismatches  255;  Indels   53;  Gaps   15;

Qy          1 MSNINV--IKSNIQAGL----------------------NSTKSGLKNLYLAIPKDYDPQ 36
Db          1 MSKITLSSLKSSLQQGLKNGKNKLNQAGTTLKNGLTQTGHSLQNGAKKLILYIPQGYDSG 60
```

```
Qy      37 KGGTLNDFIKAADELGIARLAEEPNHTETAKKSVDTVNQFLSLTQTGIAISATKLEKFLQ 96
Db      61 QGNGVQDLVKAANDLGIEVWREERSNLDIAKTSFDTTQKILGFTDRGIVLFAPQLDNLLK 120

Qy      97 KHSTNKLAKGLDSVENIDRKLGKASNVLSTLSSFLGTALAGIELDSLIKKGDAAPDALAK 156
Db     121 KNP--KIGNTLGSASSISQNIGKANTVLGGIQSILGSVLSGVNLNELLQNKDPNQLELAK 178

Qy     157 ASIDLINEIIGNLSQSTQTIEAFSSQLAKLGSTISQAKGFSNIGNKLQNL-NFSKTNLGL 215
Db     179 AGLELTNELVGNIASSVQTVDAFAEQISKLGSHLQNVKGLGGLSNKLQNLPDLGKASLGL 238

Qy     216 EIITGLLSGISAGFALADKNASTGKKVAAGFELSNQVIGNVTKAISSYVLAQRVAAGLST 275
Db     239 DIISGLLSGASAGLILADKEASTEKKAAAGVEFANQIIGNVTKAVSSYILAQRVASGLSS 298

Qy     276 TGAVAALITSSIMLAISPLAFMNAADKFNHANALDEFAKQFRKFGYDGDHLLAEYQRGVG 335
Db     299 TGPVAALIASTVALAVSPLSFLNVADKFKQADLIKSYSERFQKLGYDGDRLLADFHRETG 358

Qy     336 TIEASLTTISTALGAVSAGVSAAAVGSAVGTPIALLVAGVTGLISGILEASKQAMFESVA 395
Db     359 TIDASVTTINTALAAISGGVGAASAGSLVGAPVALLVAGVTGLITTILEYSKQAMFEHVA 418

Qy     396 NRLQGKILEWEKQNGGQNYFDKGYDSRYAAYLANNLKFLSELNKELEAERVIAITQQRWD 455
Db     419 NKVHDRIVEWEKKH-NKNYFEQGYDSRHLADLQDNMKFLINLNKELQAERVVAITQQRWD 477

Qy     456 NNIGELAGITKLGERIKSGKAYADAFEDGKKVEAGSNITLDAKTGIIDISNSNGKKTQAL 515
Db     478 NQIGDLAAISRRTDKISSGKAYVDAFEEGQHQSYDSSVQLDNKNGIINISNTN-RKTQSV 536

Qy     516 HFTSPLLTAGTESRERLTNGKYSYINKLKFGRVKNWQVTDGEASSKLDFSKVIQRVA--- 572
Db     537 LFRTPLLTPGEENRERIQEGKNSYITKLHIQRVDSWTVTDGDASSSVDFTNVVQRIAVKF 596

Qy     573 -------ETEGTDEIGLIVNAKAGNDDIFVGQGKMNIDGGDGHDRVFYSKDGGFGNITVD 625
Db     597 DDAGNIIESKDTK---IIANLGAGNDNVFVGSSTTVIDGGDGHDRVHYSR-GEYGALVID 652

Qy     626 GTSATEAGSYTVNRKVARGDIYHEVVKRQETKVGKRTETIQYRDYELRKVGYGYQSTDNL 685
Db     653 ATAETEKGSYSVKRYVGDSKALHETIATHQTNVGNREEKIEYR-REDDRFHTGYTVTDSL 711

Qy     686 KSVEEVIGSQFNDVFKGSKFNDIFHSGEGDDLLDGGAGDDRLFGGKGNDRLSGDEGDDLL 745
Db     712 KSVEEIIGSQFNDIFKGSQFDDVFHGGNGVDTIDGNDGDDHLFGGAGDDVIDGGNGNNFL 771

Qy     746 DGGSGDDVLNGGAGNDVYIFRKGDGNDTLYDGTGNDKLAFADANISDIMIERTKEGIIVK 805
Db     772 VGGTGNDIISGGKDNDIYVHKTGDGNDSITDSGGQDKLAFSDVNLKDLTFKKVDSSLEI- 830

Qy     806 RNDHSGSINIPRWY----ITSNLQNYQSNKTDHKIEQLIGKDGSYITSDQIDKILQDKKD 861
Db     831 INQKGEKVRIGNWFLEDDLASTVANYKAT-NDRKIEEIIGKGGERITSEQVDKLI--KEG 887

Qy     862 GTVITSQELKKLADENKSQKLSASDIASSLNKLVGSMALFGTANSVSSN 910
Db     888 NNQISAEALSKVVNDYNTSK-DRQNVSNSLAKLISSVGSFTSSSDFRNN 935
```

RESULT 30
AAW22156

```
ID   AAW22156 standard; Protein; 956 AA.
XX
AC   AAW22156;
XX
DT   16-FEB-1998  (first entry)
XX
DE   ApxIIC protein.
XX
KW   RTX toxin; apxICA gene; apxIBD gene; apxIIAB'C gene; apxIIIABCD gene;
KW   repeat in toxins toxin; cell-associated RTX toxin; vaccine production;
KW   therapy; A. pleuropneumoniae infection; swine pleuropneumonia.
XX
OS   Actinobacillus pleuropneumoniae.
XX
PN   CA2170839-A.
XX
PD   02-SEP-1996.
XX
PF   01-MAR-1996;   96CA-2170839.
XX
PR   01-MAR-1995;   95US-0396244.
XX
PA   (UYGU-) UNIV GUELPH.
XX
PI   MacInnes J,  Mallard B,  Ricciatti P,  Rosendal S;
XX
DR   WPI; 1997-245536/23.
DR   N-PSDB; AAT73219.
XX
PT   Preparations of microorganisms producing cell-associated RTX toxins
PT   - especially for production of vaccines against swine
PT   pleuro-pneumonia
XX
PS   Disclosure; Pages 101-104; 151pp; English.
XX
CC   AAW22151-W22161 represent A. pleuropneumoniae RTX (repeat in toxins)
CC   toxins. These sequence are encoded by the apxICA, apxIBD, apxIIAB'C, and
CC   apxIIIABCD genes (see AAT73217-T73220), and can be expressed by
CC   microorganisms used in the preparations of the invention. The
CC   preparations are bacterial preparations comprising one or more isolated
CC   and purified strains of a microorganism that produces one or more RTX
CC   toxins, where the strains have at least one cell-associated RTX toxin.
CC   The preparations are used for production of vaccines for the prophylaxis
CC   and treatment of infectious diseases caused by microorganisms that
CC   produce RTX toxins, where the strains have been attenuated or
CC   inactivated. The vaccines are preferably against Actinobacillus
CC   pleuropneumoniae infection (swine pleuropneumonia). It has been found
CC   that A. pleuropneumonia produces significant quantities of
CC   cell-associated RTX toxins when cultured under certain conditions, and
CC   that the whole-cell protein composition of the cultures corresponds to
CC   the whole-cell protein profiles obtained from cells recovered at
CC   necropsy from the pleural fluid of infected swine. Vaccination with a
CC   bacterin prepared from heat-inactivated cultures having significant
CC   quantities of cell-associated RTX toxins give significant protection of
CC   swine against challenge with homologous strains.
XX
SQ   Sequence   956 AA;

  Query Match             48.1%;  Score 2236.5;  DB 18;  Length 956;
  Best Local Similarity   47.7%;  Pred. No. 7.4e-131;
  Matches  453;  Conservative  188;  Mismatches  255;  Indels   53;  Gaps   15;

Qy       1 MSNINV--IKSNIQAGL---------------------NSTKSGLKNLYLAIPKDYDPQ 36
Db       1 MSKITLSSLKSSLQQGLKNGKNKLNQAGTTLKNGLTQTGHSLQNGAKKLILYIPQGYDSG 60

Qy      37 KGGTLNDFIKAADELGIARLAEEPNHTETAKKSVDTVNQFLSLTQTGIAISATKLEKFLQ 96
Db      61 QGNGVQDLVKAANDLGIEVWREERSNLDIAKTSFDTTQKILGFTDRGIVLFAPQLDNLLK 120

Qy      97 KHSTNKLAKGLDSVENIDRKLGKASNVLSTLSSFLGTALAGIELDSLIKKGDAAPDALAK 156
Db     121 KNP--KIGNTLGSASSISQNIGKANTVLGGIQSILGSVLSGVNLNELLQNKDPNQLELAK 178

Qy     157 ASIDLINEIIGNLSQSTQTIEAFSSQLAKLGSTISQAKGFSNIGNKLQNL-NFSKTNLGL 215
Db     179 AGLELTNELVGNIASSVQTVDAFAEQISKLGSHLQNVKGLGGLSNKLQNLPDLGKASLGL 238

Qy     216 EIITGLLSGISAGFALADKNASTGKKVAAGFELSNQVIGNVTKAISSYVLAQRVAAGLST 275
Db     239 DIISGLLSGASAGLILADKEASTEKKAAAGVEFANQIIGNVTKAVSSYILAQRVASGLSS 298

Qy     276 TGAVAALITSSIMLAISPLAFMNAADKFNHANALDEFAKQFRKFGYDGDHLLAEYQRGVG 335
```

```
                 || ||||| |:: ||:|||:|:| ||||  |: :  ::::|:| ||||| |||:: |  |
Db        299 TGPVAALIASTVALAVSPLSFLNVADKFKQADLIKSYSERFQKLGYDGDRLLADFHRETG 358

Qy        336 TIEASLTTISTALGAVSAGVSAAAVGSAVGTPIALLVAGVTGLISGILEASKQAMFESVA 395
              ||:||:|||:||| |:| || ||: || || |:||||||||||: ||| |||||| ||
Db        359 TIDASVTTINTALAAISGGVGAASAGSLVGAPVALLVAGVTGLITTILEYSKQAMFEHVA 418

Qy        396 NRLQGKILEWEKQNGGQNYFDKGYDSRYAAYLANNLKFLSELNKELEAERVIAITQQRWD 455
              |::  :|:||||::  :|||::||||: | | :|:|||  |||||:||||:||||||||
Db        419 NKVHDRIVEWEKKH-NKNYFEQGYDSRHLADLQDNMKFLINLNKELQAERVVAITQQRWD 477

Qy        456 NNIGELAGITKLGERIKSGKAYADAFEDGKKVEAGSNITLDAKTGIIDISNSNGKKTQAL 515
              | ||:|| |::  ::| ||||| ||||:|:      |:: || | |||:|||:| :|||::
Db        478 NQIGDLAAISRRTDKISSGKAYVDAFEEGQHQSYDSSVQLDNKNGIINISNTN-RKTQSV 536

Qy        516 HFTSPLLTAGTESRERLTNGKYSYINKLKFGRVKNWQVTDGEASSKLDFSKVIQRVA--- 572
               |:|||| | |:|||:  || ||| ||    || :| ||||:||| :||: |:||:|
Db        537 LFRTPLLTPGEENRERIQEGKNSYITKLHIQRVDSWTVTDGDASSSVDFTNVVQRIAVKF 596

Qy        573 -------ETEGTDEIGLIVNAKAGNDDIFVGQGKMNIDGGDGHDRVFYSKDGGFGNITVD 625
                     |:: |    :| |  ||||::|||     ||||||||| ||: | :| : :|
Db        597 DDAGNIIESKDTK---IIANLGAGNDNVFVGSSTTVIDGGDGHDRVHYSR-GEYGALVID 652

Qy        626 GTSATEAGSYTVNRKVARGDIYHEVVKRQETKVGKRTETIQYRDYELRKVGYGYQSTDNL 685
              |: || |||:| |          ||  :    :| || | | |:||  |   :   || ||:|
Db        653 ATAETEKGSYSVKRYVGDSKALHETIATHQTNVGNREEKIEYR-REDDRFHTGYTVTDSL 711

Qy        686 KSVEEVIGSQFNDVFKGSKFNDIFHSGEGDDLLDGGAGDDRLFGGKGNDRLSGDEGDDLL 745
              |||||:|||||||||:||||:|:|| | |  | || |  |||||  |||| |:|  :  |:: |
Db        712 KSVEEIIGSQFNDIFKGSQFDDVFHGGNGVDTIDGNDGDDHLFGGAGDDVIDGGNGNNFL 771

Qy        746 DGGSGDDVLNGGAGNDVYIFRKGDGNDTLYDGTGNDKLAFADANISDIMIERTKEGIIVK 805
               ||:|:|:::|| ||:|: : |||||::|  | |||||:| |: |:  ::      : :
Db        772 VGGTGNDIISGGKDNDIYVHKTGDGNDSITDSGGQDKLAFSDVNLKDLTFKKVDSSLEI- 830

Qy        806 RNDHSGSINIPRWY----ITSNLQNYQSNKTDHKIEQLIGKDGSYITSDQIDKILQDKKD 861
               |      :  |  |:   :  | : ||::   | |||::|||  |   |||:|:||:: |:
Db        831 INQKGEKVRIGNWFLEDDLASTVANYKAT-NDRKIEEIIGKGGERITSEQVDKLI--KEG 887

Qy        862 GTVITSQELKKLADENKSQKLSASDIASSLNKLVGSMALFGTANSVSSN 910
                |::: | |: :: :  :  |    ::::|| ||: |:  | :::   :|
Db        888 NNQISAEALSKVVNDYNTSK-DRQNVSNSLAKLISSVGSFTSSSDFRNN 935
```

RESULT 31
AAY51410
ID   AAY51410 standard; protein; 956 AA.
XX
AC   AAY51410;
XX
DT   05-MAY-2000 (first entry)
XX
DE   A. pleuropneumoniae clyIIA protein.
XX
KW   RTX toxin; vaccine; Repeats in Toxins; immunostimulatory; antimicrobial;
KW   antiinflammatory; antiarthritic; antiabortive; treatment; pneumonia;
KW   pleuropneumonia; septicemia; nephritis; arthritis; endocarditis;
KW   shipping fever; abortion; whooping cough; sleepy foal disease; joint ill;
KW   urinary infection; peritonitis; meningitis; gastroenteritis;
KW   passive immunization; clyIIA.
XX
OS   Actinobacillus pleuropneumoniae.
XX
PN   US6019984-A.
XX
PD   01-FEB-2000.
XX
PF   23-DEC-1996;   96US-0772270.
XX
PR   01-MAR-1995;   95US-0396244.
XX
PA   (UYGU-) UNIV GUELPH.
XX
PI   Mallard B, Rosendal S, MacInnes J, Ricciatti P;
XX
DR   WPI; 2000-146864/13.
DR   N-PSDB; AAZ88586.
XX
PT   Bacterial preparation comprising microorganisms which produce a member
PT   of the Repeats in Toxins (RTX) family, useful for treating swine
PT   pleuropneumonia, arthritis in swine, shipping fever and abortion in
PT   cattle, and sleepy foal disease -
XX
PS   Disclosure; Column 67-72; 96pp; English.
XX
CC   This invention describes a novel bacterial preparation (I) which
CC   comprises one or more isolated and purified strain(s) of a microorganism,
CC   cultured in tryptone yeast extract (TYE) broth, which produces one or
CC   more RTX toxins (belonging to the family of toxins referred to as Repeats
CC   in Toxins), where the strain(s) have at least one RTX toxin which is
CC   cell-associated. The products of the invention have immunostimulatory,
CC   antimicrobial, antiinflammatory, antiarthritic and antiabortive activity.
CC   The bacterial preparation may be used as vaccines for the prophylaxis and
CC   treatment of infectious diseases caused by strains of microorganisms
CC   which produce one or more RTX toxins. The infectious diseases are swine
CC   pleuropneumonia, pneumonia, septicemia, nephritis and arthritis in swine;
CC   septicemia, nephritis, endocarditis and arthritis in piglets; shipping
CC   fever and abortion in cattle; whooping cough, sleepy foal disease or
CC   joint ill (purulent nephritis, arthritis) in foals; septicemia,
CC   polyarthritis and abortion in horses; and urinary infections,
CC   peritonitis, meningitis, and gastroenteritis. The bacterial preparations
CC   may also be used to prepare antibodies which may be used as a means of
CC   passive immunization. This sequence represents the Actinobacillus
CC   pleuropneumoniae clyIIA protein described in the method of the invention.
XX
SQ   Sequence   956 AA;

```
  Query Match             48.1%;  Score 2236.5;  DB 21;  Length 956;
  Best Local Similarity   47.7%;  Pred. No. 7.4e-131;
  Matches  453;  Conservative  188;  Mismatches  255;  Indels   53;  Gaps   15;

Qy          1 MSNINV--IKSNIQAGL---------------------NSTKSGLKNLYLAIPKDYDPQ 36
              || |  :  :||::| ||                    :| ::| | | | ||: ||
Db          1 MSKITLSSLKSSLQQGLKNGKNKLNQAGTTLKNGLTQTGHSLQNGAKKLILYIPQGYDSG 60

Qy         37 KGGTLNDFIKAADELGIARLAEEPNHTETAKKSVDTVNQFLSLTQTGIAISATKLEKFLQ 96
              :|  :  | :|||::|||     || | :: : || | ||   :  |   |  || : | :|:  |:
Db         61 QGNGVQDLVKAANDLGIEVWREERSNLDIAKTSFDTTQKILGFTDRGIVLFAPQLDNLLK 120

Qy         97 KHSTNKLAKGLDSVENIDRKLGKASNVLSTLSSFLGTALAGIELDSLIKKGDAAPDALAK 156
              |:   |:   |  |   :  :  :|||: ||  |  || |: |:|: |: |:  |     |||
Db        121 KNP--KIGNTLGSASSISQNIGKANTVLGGIQSILGSVLSGVNLNELLQNKDPNQLELAK 178

Qy        157 ASIDLINEIIGNLSQSTQTIEAFSSQLAKLGSTISQAKGFSNIGNKLQNL-NFSKTNLGL 215
              | ::| ||::||:: | ||::||: |::|||| :   ||    : ||||| :  | :|||
Db        179 AGLELTNELVGNIASSVQTVDAFAEQISKLGSHLQNVKGLGGLSNKLQNLPDLGKASLGL 238

Qy        216 EIITGLLSGISAGFALADKNASTGKKVAAGFELSNQVIGNVTKAISSYVLAQRVAAGLST 275
              :||:||||| |||  |||| ||| || ||| | :||:||||||:|||:||||||:|||:
Db        239 DIISGLLSGASAGLILADKEASTEKKAAAGVEFANQIIGNVTKAVSSYILAQRVASGLSS 298

Qy        276 TGAVAALITSSIMLAISPLAFMNAADKFNHANALDEFAKQFRKFGYDGDHLLAEYQRGVG 335
              || ||||| |:: ||:|||:|:| ||||  |: :  ::::|:| ||||| |||:: |  |
Db        299 TGPVAALIASTVALAVSPLSFLNVADKFKQADLIKSYSERFQKLGYDGDRLLADFHRETG 358

Qy        336 TIEASLTTISTALGAVSAGVSAAAVGSAVGTPIALLVAGVTGLISGILEASKQAMFESVA 395
              ||:||:|||:||| |:| || ||: || || |:||||||||||: ||| |||||| ||
Db        359 TIDASVTTINTALAAISGGVGAASAGSLVGAPVALLVAGVTGLITTILEYSKQAMFEHVA 418

Qy        396 NRLQGKILEWEKQNGGQNYFDKGYDSRYAAYLANNLKFLSELNKELEAERVIAITQQRWD 455
              |::  :|:||||::  :|||::||||: | | :|:|||  |||||:||||:||||||||
Db        419 NKVHDRIVEWEKKH-NKNYFEQGYDSRHLADLQDNMKFLINLNKELQAERVVAITQQRWD 477

Qy        456 NNIGELAGITKLGERIKSGKAYADAFEDGKKVEAGSNITLDAKTGIIDISNSNGKKTQAL 515
              | ||:|| |::  ::| ||||| ||||:|:      |:: || | |||:|||:| :|||::
Db        478 NQIGDLAAISRRTDKISSGKAYVDAFEEGQHQSYDSSVQLDNKNGIINISNTN-RKTQSV 536
```

```
Qy        516 HFTSPLLTAGTESRERLTNGKYSYINKLKFGRVKNWQVTDGEASSKLDFSKVIQRVA--- 572
Db        537 LFRTPLLTPGEENRERIQEGKNSYITKLHIQRVDSWTVTDGDASSSVDFTNVVQRIAVKF 596

Qy        573 -------ETEGTDEIGLIVNAKAGNDDIFVGQGKMNIDGGDGHDRVFYSKDGGFGNITVD 625
Db        597 DDAGNIIESKDTK---IIANLGAGNDNVFVGSSTTVIDGGDGHDRVHYSR-GEYGALVID 652

Qy        626 GTSATEAGSYTVNRKVARGDIYHEVVKRQETKVGKRTETIQYRDYELRKVGYGYQSTDNL 685
Db        653 ATAETEKGSYSVKRYVGDSKALHETIATHQTNVGNREEKIEYR-REDDRFHTGYTVTDSL 711

Qy        686 KSVEEVIGSQFNDVFKGSKFNDIFHSGEGDDLLDGGAGDDRLFGGKGNDRLSGDEGDDLL 745
Db        712 KSVEEIIGSQFNDIFKGSQFDDVFHGGNGVDTIDGNDGDDHLFGGAGDDVIDGGNGNNFL 771

Qy        746 DGGSGDDVLNGGAGNDVYIFRKGDGNDTLYDGTGNDKLAFADANISDIMIERTKEGIIVK 805
Db        772 VGGTGNDIISGGKDNDIYVHKTGDGNDSITDSGGQDKLAFSDVNLKDLTFKKVDSSLEI- 830

Qy        806 RNDHSGSINIPRWY----ITSNLQNYQSNKTDHKIEQLIGKDGSYITSDQIDKILQDKKD 861
Db        831 INQKGEKVRIGNWFLEDDLASTVANYKAT-NDRKIEEIIGKGGERITSEQVDKLI--KEG 887

Qy        862 GTVITSQELKKLADENKSQKLSASDIASSLNKLVGSMALFGTANSVSSN 910
Db        888 NNQISAEALSKVVNDYNTSK-DRQNVSNSLAKLISSVGSFTSSSDFRNN 935
```

RESULT 32
AAW22159
ID   AAW22159 standard; Protein; 1049 AA.
XX
AC   AAW22159;
XX
DT   16-FEB-1998 (first entry)
XX
DE   ApxIIIB protein.
XX
KW   RTX toxin; apxICA gene; apxIBD gene; apxIIAB'C gene; apxIIIABCD gene;
KW   repeat in toxins toxin; cell-associated RTX toxin; vaccine production;
KW   therapy; A. pleuropneumoniae infection; swine pleuropneumonia.
XX
OS   Actinobacillus pleuropneumoniae.
XX
PN   CA2170839-A.
XX
PD   02-SEP-1996.
XX
PF   01-MAR-1996;   96CA-2170839.
XX
PR   01-MAR-1995;   95US-0396244.
XX
PA   (UYGU-) UNIV GUELPH.
XX
PI   MacInnes J, Mallard B, Ricciatti P, Rosendal S;
XX
DR   WPI; 1997-245536/23.
DR   N-PSDB; AAT73220.
XX
PT   Preparations of microorganisms producing cell-associated RTX toxins
PT   - especially for production of vaccines against swine
PT   pleuro-pneumonia
XX
PS   Disclosure; Pages 107-110; 151pp; English.
XX
CC   AAW22151-W22161 represent A. pleuropneumoniae RTX (repeat in toxins)
CC   toxins. These sequence are encoded by the apxICA, apxIBD, apxIIAB'C, and
CC   apxIIIABCD genes (see AAT73217-T73220), and can be expressed by
CC   microorganisms used in the preparations of the invention. The
CC   preparations are bacterial preparations comprising one or more isolated
CC   and purified strains of a microorganism that produces one or more RTX
CC   toxins, where the strains have at least one cell-associated RTX toxin.
CC   The preparations are used for production of vaccines for the prophylaxis
CC   and treatment of infectious diseases caused by microorganisms that
CC   produce RTX toxins, where the strains have been attenuated or
CC   inactivated. The vaccines are preferably against Actinobacillus
CC   pleuropneumoniae infection (swine pleuropneumonia). It has been found
CC   that A. pleuropneumonia produces significant quantities of
CC   cell-associated RTX toxins when cultured under certain conditions, and
CC   that the whole-cell protein composition of the cultures corresponds to
CC   the whole-cell protein profiles obtained from cells recovered at
CC   necropsy from the pleural fluid of infected swine. Vaccination with a
CC   bacterin prepared from heat-inactivated cultures having significant
CC   quantities of cell-associated RTX toxins give significant protection of
CC   swine against challenge with homologous strains.
XX
SQ   Sequence   1049 AA;

  Query Match   41.4%; Score 1924; DB 18; Length 1049;
  Best Local Similarity   43.0%; Pred. No. 2.5e-111;
  Matches 428; Conservative 152; Mismatches 296; Indels 120; Gaps   21;

```
Qy          8 KSNIQAGLNSTKSGLK---------------------NLYLAIPKDYDPQKGGTLNDFI 45
Db         19 KRQVKKGYDVTKNGLQYGVSQAKLQALAAGKAVQKYGNKLVLVIPKEYDGSVGNGFFDLV 78

Qy         46 KAADELGIARLAEEPNHTETAKKSVDTVNQFLSLTQTGIAISATKLEKFLQKHSTNKLAK 105
Db         79 KAAEELGIQVKYVNRNELEVAHKSLGTADQFLGLTERGLTLFAPQLDQFLQKHSKISNVV 138

Qy        106 GLDSVENIDRKLGKASNVLSTLSSFLGTALAGIELDSLIKKGDAAPDALAKASIDLINEI 165
Db        139 GSSTGDAVS-KLAKSQTIISGIQSVLGTVLAGINLNEAIISGGSELE-LAEAGVSLASEL 196

Qy        166 IGNLSQSTQTIEAFSSQLAKLGSTISQAKGFSNIGNKLQNLN---FSKTNLGLEIITGLL 222
Db        197 VSNIAKGTTTIDAFTTQIQNFGKLAENAKGLGGVGRQLQNISGSALSKTGLGLDIISSLL 256

Qy        223 SGISAGFALADKNASTGKKVAAGFELSNQVIGNVTKAISSYVLAQRVAAGLSTTGAVAAL 282
Db        257 SGVTRSFALRNKNASTSTKVAAGFELSNQVIGGITKAVSSYILAQRLRAGLSTTGPAAAL 316

Qy        283 ITSSIMLAISPLAFMNAADKFNHANALDEFAKQFRKFGYDGDHLLAEYQRGVGTIEASLT 342
Db        317 IASSISLAISPLAFLRVADNFNRSKEIGEFAERFKKLGYDGDKLLSEFYHEAGTIDASIT 376

Qy        343 TISTALGAVSAGVSAAAVGSAVGTPIALLVAGVTGLISGILEASKQAMFESVANRLQGKI 402
Db        377 TISTALSAIAAGTAAASAGALVGAPITLLVTGITGLISGILEFSKQPMLDHVASKIGNKI 436

Qy        403 LEWEKQNGGQNYFDKGYDSRYAAYLANNLKFLSELNKELEAERVIAITQQRWDNNIGELA 462
Db        437 DEWEKKY-GKNYFENGYDARHKAFLEDSFSLLSSFNKQYETERAVLITQQRWDEYIGELA 495

Qy        463 GITKLGERIKSGKAYADAFEDGKKVEAG----SNITLDAKTGIIDISNSNGKKTQALHFT 518
Db        496 GITGKGDKLSSGKAYVDYFQEGKLLEKKPDDFSKVVFDPTKGEIDISNS--QTSTLLKFV 553

Qy        519 SPLLTAGTESRERLTNGKYSYINKLKFGRVKNWQVTDG--EASSKLDFSKVIQR------ 570
Db        554 TPLLTPGTESRERTQTGKYEYITKLVVKGKDKW-VVNGVKDKGAVYDYTNLIQHAHISSS 612

Qy        571 VAETEGTDEIGLIVNAKAGNDDIFVGQGKMNIDGGDGHDRVFYSKDGGFGNITVDGTSAT 630
Db        613 VARGEEYREVRLVSHLGNGNDKVFLAAGSAEIHAGEGHDVVYYDKT-DTGLLVIDGTKAT 671

Qy        631 EAGSYTVNRKVARG-DIYHEVVKRQETKVGKRTETIQYRDYELRKVG-YGYQSTDNLKSV 688
Db        672 EQGRYSVTRELSGATKILREVIKNQKYAVGKREETLEYRDYELTQSGNSNLKAHDELHSV 731

Qy        689 EEVIGSQFNDVFKGSKFNDIFHSGEGDDLLDGGAGDDRLFGGKGNDRLSGD--------- 739
Db        732 EE-IGSNQRDEFKGSKFRDIFHGADGDDLLNGNDGDDILYGDKGNDELRGDNGNDQLYGG 790
```

```
Qy        740 EGD---------------------------------------------DLLDGGSGDDV 753
              |||                                            ||||||  |:|
Db        791 EGDDKLLGGNGNNYLSGGDGNDELQVLGNGFNVLRGGKGDDKLYGSSGSDLLDGGEGNDY 850

Qy        754 LNGGAGNDVYIFRKGDGNDTLYD---GTGNDKLAFADANISDIMIERTKEGIIVK--RND 808
              | || |:| |::|    || |:||      : :|||  :| :  :|:::|  : :  :   |
Db        851 LEGGDGSDFYVYRSTSGNHTIYDQGKASDSDKLYLSDLSFDNILVKRVNDNLEFRSNNNS 910

Qy        809 HSGSINIPRWYITSNLQNYQSNKTDHKIEQLIGKDGSYITSDQIDKILQDKKDGTVITSQ 868
              :||  :  |  |:    |     |       |||||:: |:|  :|:  :       |    |    |
Db        911 NSGVLTIKDWFKGGNSYN-------HKIEQIVDKNGRKLTAGNLGNNFHD----TQQASS 959

Qy        869 ELKKLADENKSQKLSASDIASSLNKLVGSMALFGTA 904
              || :  |       ||:  : : | |:: :    || |
Db        960 LLKNVTQEQNESNLSS--LKTELGKIITNAGNFGVA 993
```

RESULT 33
AAY51412

```
ID   AAY51412 standard; protein; 1049 AA.
XX
AC   AAY51412;
XX
DT   05-MAY-2000  (first entry)
XX
DE   A. pleuropneumoniae apxIIIA protein.
XX
KW   RTX toxin; vaccine; Repeats in Toxins; immunostimulatory; antimicrobial;
KW   antiinflammatory; antiarthritic; antiabortive; treatment; pneumonia;
KW   pleuropneumonia; septicemia; nephritis; arthritis; endocarditis;
KW   shipping fever; abortion; whooping cough; sleepy foal disease; joint ill;
KW   urinary infection; peritonitis; meningitis; gastroenteritis;
KW   passive immunization; apxIIIA.
XX
OS   Actinobacillus pleuropneumoniae.
XX
PN   US6019984-A.
XX
PD   01-FEB-2000.
XX
PF   23-DEC-1996;   96US-0772270.
XX
PR   01-MAR-1995;   95US-0396244.
XX
PA   (UYGU-) UNIV GUELPH.
XX
PI   Mallard B,  Rosendal S,  MacInnes J,  Ricciatti P;
XX
DR   WPI; 2000-146864/13.
DR   N-PSDB; AAZ88587.
XX
PT   Bacterial preparation comprising microorganisms which produce a member
PT   of the Repeats in Toxins (RTX) family, useful for treating swine
PT   pleuropneumonia, arthritis in swine, shipping fever and abortion in
PT   cattle, and sleepy foal disease -
XX
PS   Disclosure; Column 77-84; 96pp; English.
XX
CC   This invention describes a novel bacterial preparation (I) which
CC   comprises one or more isolated and purified strain(s) of a microorganism,
CC   cultured in tryptone yeast extract (TYE) broth, which produces one or
CC   more RTX toxins (belonging to the family of toxins referred to as Repeats
CC   in Toxins), where the strain(s) have at least one RTX toxin which is
CC   cell-associated. The products of the invention have immunostimulatory,
CC   antimicrobial, antiinflammatory, antiarthritic and antiabortive activity.
CC   The bacterial preparation may be used as vaccines for the prophylaxis and
CC   treatment of infectious diseases caused by strains of microorganisms
CC   which produce one or more RTX toxins. The infectious diseases are swine
CC   pleuropneumonia, pneumonia, septicemia, nephritis and arthritis in swine;
CC   septicemia, nephritis, endocarditis and arthritis in piglets; shipping
CC   fever and abortion in cattle; whooping cough, sleepy foal disease or
CC   joint ill (purulent nephritis, arthritis) in foals; septicemia,
CC   polyarthritis and abortion in horses; and urinary infections,
CC   peritonitis, meningitis, and gastroenteritis. The bacterial preparations
CC   may also be used to prepare antibodies which may be used as a means of
CC   passive immunization. This sequence represents the Actinobacillus
CC   pleuropneumoniae apxIIIA protein described in the method of the
CC   invention.
XX
SQ   Sequence   1049 AA;

  Query Match             41.4%;  Score 1924;  DB 21;  Length 1049;
  Best Local Similarity   43.0%;  Pred. No. 2.5e-111;
  Matches  428;  Conservative  152;  Mismatches  296;  Indels  120;  Gaps   21;

Qy          8 KSNIQAGLNSTKSGLK---------------------NLYLAIPKDYDPQKGGTLNDFI 45
              |  ::  | : ||:||:                     | | |||:||    |    | :
Db         19 KRQVKKGYDVTKNGLQYGVSQAKLQALAAGKAVQKYGNKLVLVIPKEYDGSVGNGFFDLV 78

Qy         46 KAADELGIARLAEEPNHTETAKKSVDTVNQFLSLTQTGIAISATKLEKFLQKHSTNKLAK 105
              |||:||||          |   |  |  ||:  | :|||  ||: |: :  |  :|::|||||
Db         79 KAAEELGIQVKYVNRNELEVAHKSLGTADQFLGLTERGLTLFAPQLDQFLQKHSKISNVV 138

Qy        106 GLDSVENIDRKLGKASNVLSTLSSFLGTALAGIELDSLIKKGDAAPDALAKASIDLINEI 165
              |  : : :   || |:   ::| :  | |||  |||| |    |   :  :  ||:|  : |:|:
Db        139 GSSTGDAVS-KLAKSQTIISGIQSVLGTVLAGINLNEAIISGGSELE-LAEAGVSLASEL 196

Qy        166 IGNLSQSTQTIEAFSSQLAKLGSTISQAKGFSNIGNKLQNLN---FSKTNLGLEIITGLL 222
              : |:::  | ||:|::|:      |      |||    :|  :|||::      ||| |||:||: ||
Db        197 VSNIAKGTTTIDAFTTQIQNFGKLAENAKGLGGVGRQLQNISGSALSKTGLGLDIISSLL 256

Qy        223 SGISAGFALADKNASTGKKVAAGFELSNQVIGNVTKAISSYVLAQRVAAGLSTTGAVAAL 282
              ||::  |||  :|||||  ||||||||||||  :|||:|||:||||: ||||||  |||
Db        257 SGVTRSFALRNKNASTSTKVAAGFELSNQVIGGITKAVSSYILAQRLRAGLSTTGPAAAL 316

Qy        283 ITSSIMLAISPLAFMNAADKFNHANALDEFAKQFRKFGYDGDHLLAEYQRGVGTIEASLT 342
              | ||| ||||||||:  || ||  :  : |||::|:| ||||| ||:|:     |||:||:|
Db        317 IASSISLAISPLAFLRVADNFNRSKEIGEFAERFKKLGYDGDKLLSEFYHEAGTIDASIT 376

Qy        343 TISTALGAVSAGVSAAAVGSAVGTPIALLVAGVTGLISGILEASKQAMFESVANRLQGKI 402
              |||||| |::|| :||:  |: || || || |||  |:|||||||| |||  |  : ||:::  ||
Db        377 TISTALSAIAAGTAAASAGALVGAPITLLVTGITGLISGILEFSKQPMLDHVASKIGNKI 436

Qy        403 LEWEKQNGGQNYFDKGYDSRYAAYLANNLKFLSELNKELEAERVIAITQQRWDNNIGELA 462
              ||||:   |:|||: |||:|: |:|: |:  ::       ||  ||: | || :  |||||||  |||||
Db        437 DEWEKKY-GKNYFENGYDARHKAFLEDSFSLLSSFNKQYETERAVLITQQRWDEYIGELA 495

Qy        463 GITKLGERIKSGKAYADAFEDGKKVEAG----SNITLDAKTGIIDISNSNGKKTQALHFT 518
              |||  |:::: ||||| | |::|| :|        | :    |  | ||||||   : :  | |
Db        496 GITGKGDKLSSGKAYVDYFQEGKLLEKKPDDFSKVVFDPTKGEIDISNS--QTSTLLKFV 553

Qy        519 SPLLTAGTESRERLTNGKYSYINKLKFGRVKNWQVTDG--EASSKLDFSKVIQR------ 570
              :|||| ||||||||    ||| || ||       | |  :|   :    :    |:: :||
Db        554 TPLLTPGTESRERTQTGKYEYITKLVVKGKDKW-VVNGVKDKGAVYDYTNLIQHAHISSS 612

Qy        571 VAETEGTDEIGLIVNAKAGNDDIFVGQGKMNIDGGDGHDRVFYSKDGGFGNITVDGTSAT 630
              || |    |: |: :    ||| :|:  |    |  |:||| |:| |      | : :||| ||
Db        613 VARGEEYREVRLVSHLGNGNDKVFLAAGSAEIHAGEGHDVVYYDKT-DTGLLVIDGTKAT 671

Qy        631 EAGSYTVNRKVARG-DIYHEVVKRQETKVGKRTETIQYRDYELRKVG-YGYQSTDNLKSV 688
              | | |:| |:::     |   ||:| |:  ||| ||::||||||| |     |    | | ||
Db        672 EQGRYSVTRELSGATKILREVIKNQKYAVGKREETLEYRDYELTQSGNSNLKAHDELHSV 731

Qy        689 EEVIGSQFNDVFKGSKFNDIFHSGEGDDLLDGGAGDDRLFGGKGNDRLSGD--------- 739
              || |||    | |||||| ||||    :|||||:|  ||| |:| |||| | ||
Db        732 EE-IGSNQRDEFKGSKFRDIFHGADGDDLLNGNDGDDILYGDKGNDELRGDNGNDQLYGG 790

Qy        740 EGD---------------------------------------------DLLDGGSGDDV 753
              |||                                             ||||||  |:|
Db        791 EGDDKLLGGNGNNYLSGGDGNDELQVLGNGFNVLRGGKGDDKLYGSSGSDLLDGGEGNDY 850

Qy        754 LNGGAGNDVYIFRKGDGNDTLYD---GTGNDKLAFADANISDIMIERTKEGIIVK--RND 808
              | || |:| |::|    || |:||      : :|||  :| :  :|:::|  : :  :   |
Db        851 LEGGDGSDFYVYRSTSGNHTIYDQGKASDSDKLYLSDLSFDNILVKRVNDNLEFRSNNNS 910

Qy        809 HSGSINIPRWYITSNLQNYQSNKTDHKIEQLIGKDGSYITSDQIDKILQDKKDGTVITSQ 868
```

```
                 :|| : |  |:    |  |        |||||:: |:|  :|:  :        |     |   |
Db        911 NSGVLTIKDWFKGGNSYN-------HKIEQIVDKNGRKLTAGNLGNNFHD----TQQASS 959

Qy        869 ELKKLADENKSQKLSASDIASSLNKLVGSMALFGTA 904
              || :  |      ||:  : : | |:: :   || |
Db        960 LLKNVTQEQNESNLSS--LKTELGKIITNAGNFGVA 993


RESULT 34
AAR54781
ID   AAR54781 standard; Protein; 1244 AA.
XX
AC   AAR54781;
XX
DT   25-MAR-2003  (updated)
DT   15-OCT-1994  (first entry)
XX
DE   Leukotoxin AppIIIA.
XX
KW   Leukotoxin; AppIIIA; pleuropneumonia; vaccine; diagnostic.
XX
OS   Actinobacillus pleuropneumoniae.
XX
FH   Key             Location/Qualifiers
FT   Domain          327..345
FT                   /label= transmembrane domain
FT   Domain          488..503
FT                   /label= transmembrane domain
FT   Domain          570..587
FT                   /label= transmembrane domain
XX
PN   WO9409821-A1.
XX
PD   11-MAY-1994.
XX
PF   04-NOV-1993;   93WO-US10500.
XX
PR   05-NOV-1992;   92US-0972229.
PR   03-JUN-1993;   93US-0072285.
XX
PA   (CORR ) CORNELL RES FOUND INC.
XX
PI   Chang Y;
XX
DR   WPI; 1994-167130/20.
DR   N-PSDB; AAQ64827.
XX
PT   DNA encoding Actinobacillus pleuropneumoniae leukotoxin - used to
PT   prepare prods. for use in vaccines for porcine pleuropneumonia
PT   and as diagnostic reagents
XX
PS   Disclosure; Fig 3; 65pp; English.
XX
CC   The AppIIIA gene having the sequence given in AAQ64827 was isolated
CC   from phage lambda clones yfc 26-28 and yfc 31-32 of a genomic
CC   library of A. pleuropneumoniae ser. 2.  The gene encodes a novel
CC   leukotoxin (sequence AAR54781).
CC   (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ   Sequence   1244 AA;

  Query Match             41.4%;  Score 1924;  DB 15;  Length 1244;
  Best Local Similarity   43.0%;  Pred. No. 3.1e-111;
  Matches  428;  Conservative  152;  Mismatches  296;  Indels  120;  Gaps   21;

Qy          8 KSNIQAGLNSTKSGLK---------------------NLYLAIPKDYDPQKGGTLNDFI 45
              |  :: | : ||:||:                     | | |||:||   |     | :
Db        192 KRQVKKGYDVTKNGLQYGVSQAKLQALAAGKAVQKYGNKLVLVIPKEYDGSVGNGFFDLV 251

Qy         46 KAADELGIARLAEEPNHTETAKKSVDTVNQFLSLTQTGIAISATKLEKFLQKHSTNKLAK 105
              |||:||||          |  | | ||: | :|||  ||: |: :  |  :|::||||||
Db        252 KAAEELGIQVKYVNRNELEVAHKSLGTADQFLGLTERGLTLFAPQLDQFLQKHSKISNVV 311

Qy        106 GLDSVENIDRKLGKASNVLSTLSSFLGTALAGIELDSLIKKGDAAPDALAKASIDLINEI 165
              |  : : :  || |:  ::| : | ||| |||| |:  |   |  :   : ||:| : | :|:
Db        312 GSSTGDAVS-KLAKSQTIISGIQSVLGTVLAGINLNEAIISGGSELE-LAEAGVSLASEL 369

Qy        166 IGNLSQSTQTIEAFSSQLAKLGSTISQAKGFSNIGNKLQNLN---FSKTNLGLEIITGLL 222
              : |::: | ||:||::|:  |          |||   :| :|||::      ||| |||:||: ||
Db        370 VSNIAKGTTTIDAFTTQIQNFGKLAENAKGLGGVGRQLQNISGSALSKTGLGLDIISSLL 429

Qy        223 SGISAGFALADKNASTGKKVAAGFELSNQVIGNVTKAISSYVLAQRVAAGLSTTGAVAAL 282
              ||::  ||| :|||||  |||||||||||||| :|||:|||:||||: |||||||  |||
Db        430 SGVTRSFALRNKNASTSTKVAAGFELSNQVIGGITKAVSSYILAQRLRAGLSTTGPAAAL 489

Qy        283 ITSSIMLAISPLAFMNAADKFNHANALDEFAKQFRKFGYDGDHLLAEYQRGVGTIEASLT 342
              | ||| ||||||||:  || ||  :  :  |||::|:| ||||| ||:|:     |||:||:|
Db        490 IASSISLAISPLAFLRVADNFNRSKEIGEFAERFKKLGYDGDKLLSEFYHEAGTIDASIT 549

Qy        343 TISTALGAVSAGVSAAAVGSAVGTPIALLVAGVTGLISGILEASKQAMFESVANRLQGKI 402
              ||||||  |::|| :||: |: || || ||| |:||||||||| ||| | : ||::: ||
Db        550 TISTALSAIAAGTAAASAGALVGAPITLLVTGITGLISGILEFSKQPMLDHVASKIGNKI 609

Qy        403 LEWEKQNGGQNYFDKGYDSRYAAYLANNLKFLSELNKELEAERVIAITQQRWDNNIGELA 462
              ||||:  |:|||: |||:|: |:| ::   ||  ||: | || : |||||||  |||||
Db        610 DEWEKKY-GKNYFENGYDARHKAFLEDSFSLLSSFNKQYETERAVLITQQRWDEYIGELA 668

Qy        463 GITKLGERIKSGKAYADAFEDGKKVEAG----SNITLDAKTGIIDISNSNGKKTQALHFT 518
              |||  |:::  ||||  |::||| :|        |  |  ||||||  :  :    ||
Db        669 GITGKGDKLSSGKAYVDYFQEGKLLEKKPDDFSKVVFDPTKGEIDISNS--QTSTLLKFV 726

Qy        519 SPLLTAGTESRERLTNGKYSYINKLKFGRVKNWQVTDG--EASSKLDFSKVIQR------ 570
              :|||| |||||||  |||  || ||        |  :|   :   :  |:: :||
Db        727 TPLLTPGTESRERTQTGKYEYITKLVVKGKDKW-VVNGVKDKGAVYDYTNLIQHAHISSS 785

Qy        571 VAETEGTDEIGLIVNAKAGNDDIFVGQGKMNIDGGDGHDRVFYSKDGGFGNITVDGTSAT 630
              ||  |    |:  |: :    ||| :|:   |    |   |:||| |:|  |     | : :||| ||
Db        786 VARGEEYREVRLVSHLGNGNDKVFLAAGSAEIHAGEGHDVVYYDKT-DTGLLVIDGTKAT 844

Qy        631 EAGSYTVNRKVARG-DIYHEVVKRQETKVGKRTETIQYRDYELRKVG-YGYQSTDNLKSV 688
              | | |:| |:::     |  ||:| |:  ||||  ||::|||||| :  |     :: | | ||
Db        845 EQGRYSVTRELSGATKILREVIKNQKYAVGKREETLEYRDYELTQSGNSNLKAHDELHSV 904

Qy        689 EEVIGSQFNDVFKGSKFNDIFHSGEGDDLLDGGAGDDRLFGGKGNDRLSGD--------- 739
              || |||    | |||||| ||||  | :||||:|  ||| |:| |||| ||
Db        905 EE-IGSNQRDEFKGSKFRDIFHGADGDDLLNGNDGDDILYGDKGNDELRGDNGNDQLYGG 963

Qy        740 EGD------------------------------------------------DLLDGGSGDDV 753
              |||                                                |||||| |:|
Db        964 EGDDKLLGGNGNNYLSGGDGNDELQVLGNGFNVLRGGKGDDKLYGSSGSDLLDGGEGNDY 1023

Qy        754 LNGGAGNDVYIFRKGDGNDTLYD---GTGNDKLAFADANISDIMIERTKEGIIVK--RND 808
              | || |:| |::|    || |:||      : :|||   :| :  :|:::|   : :  :    |
Db       1024 LEGGDGSDFYVYRSTSGNHTIYDQGKASDSDKLYLSDLSFDNILVKRVNDNLEFRSNNNS 1083

Qy        809 HSGSINIPRWYITSNLQNYQSNKTDHKIEQLIGKDGSYITSDQIDKILQDKKDGTVITSQ 868
              :|| :  |  |:    |  |        |||||:: |:|  :|:  :        |     |   |
Db       1084 NSGVLTIKDWFKGGNSYN-------HKIEQIVDKNGRKLTAGNLGNNFHD----TQQASS 1132

Qy        869 ELKKLADENKSQKLSASDIASSLNKLVGSMALFGTA 904
              || :  |      ||:  : : | |:: :   || |
Db       1133 LLKNVTQEQNESNLSS--LKTELGKIITNAGNFGVA 1166


RESULT 35
AAW22152
ID   AAW22152 standard; Protein; 1022 AA.
XX
AC   AAW22152;
XX
DT   16-FEB-1998  (first entry)
XX
DE   ApxIA protein.
XX
KW   RTX toxin; apxICA gene; apxIBD gene; apxIIAB'C gene; apxIIIABCD gene;
KW   repeat in toxins toxin; cell-associated RTX toxin; vaccine production;
```

```
KW   therapy; A. pleuropneumoniae infection; swine pleuropneumonia.
XX
OS   Actinobacillus pleuropneumoniae.
XX
PN   CA2170839-A.
XX
PD   02-SEP-1996.
XX
PF   01-MAR-1996;    96CA-2170839.
XX
PR   01-MAR-1995;    95US-0396244.
XX
PA   (UYGU-) UNIV GUELPH.
XX
PI   MacInnes J,  Mallard B,  Ricciatti P,  Rosendal S;
XX
DR   WPI; 1997-245536/23.
DR   N-PSDB; AAT73217.
XX
PT   Preparations of microorganisms producing cell-associated RTX toxins
PT   - especially for production of vaccines against swine
PT   pleuro-pneumonia
XX
PS   Disclosure; Pages 88-91; 151pp; English.
XX
CC   AAW22151-W22161 represent A. pleuropneumoniae RTX (repeat in toxins)
CC   toxins. These sequence are encoded by the apxICA, apxIBD, apxIIAB'C, and
CC   apxIIIABCD genes (see AAT73217-T73220), and can be expressed by
CC   microorganisms used in the preparations of the invention. The
CC   preparations are bacterial preparations comprising one or more isolated
CC   and purified strains of a microorganism that produces one or more RTX
CC   toxins, where the strains have at least one cell-associated RTX toxin.
CC   The preparations are used for production of vaccines for the prophylaxis
CC   and treatment of infectious diseases caused by microorganisms that
CC   produce RTX toxins, where the strains have been attenuated or
CC   inactivated. The vaccines are preferably against Actinobacillus
CC   pleuropneumoniae infection (swine pleuropneumonia). It has been found
CC   that A. pleuropneumonia produces significant quantities of
CC   cell-associated RTX toxins when cultured under certain conditions, and
CC   that the whole-cell protein composition of the cultures corresponds to
CC   the whole-cell protein profiles obtained from cells recovered at
CC   necropsy from the pleural fluid of infected swine. Vaccination with a
CC   bacterin prepared from heat-inactivated cultures having significant
CC   quantities of cell-associated RTX toxins give significant protection of
CC   swine against challenge with homologous strains.
XX
SQ   Sequence   1022 AA;

  Query Match             37.9%;  Score 1763.5;  DB 18;  Length 1022;
  Best Local Similarity   40.0%;  Pred. No. 2.4e-101;
  Matches  400;  Conservative  181;  Mismatches  309;  Indels  109;  Gaps   26;

Qy          8 KSNIQAGLNSTKSGL-------KNLYLAIPKDYDPQKGGTLNDFIKAADELGIARLAEEP 60
                ||  ::|  : |:||        : | | |||||   | :||| :|||: |||     |
Db         21 KSAAKSGAGALKNGLGQVKQAGQKLILYIPKDYQASTGSSLNDLVKAAEALGIEVHRSEK 80

Qy         61 NHTETAKKSVDTVNQFLSLTQTGIAISATKLEKFLQKHSTNKLAKGL-DSVENIDRKLGK 119
              | |  ||:          |  :  |  :: |||:  | : :| |  |:  |   | : ::| |
Db         81 NGTALAKELFGTTEKLLGFSERGIALFAPQFDKLLNKN--QKLSKSLGGSSEALGQRLNK 138

Qy        120 ASNVLSTLSSFLGTALAGIELDSLIKK----GDAAPDALAKASIDLINEIIGNLSQSTQT 175
                   || |  ||||||:||::||||:::          |  :  ||||  :||   ::: |:: :|
Db        139 TQTALSALQSFLGTAIAGMDLDSLLRRRRNGEDVSGSELAKAGVDLAAQLVDNIASATGT 198

Qy        176 IEAFSSQLAKLGSTISQAKGFSNIGNKLQNL-NFSKTNLGLEIITGLLSGISAGFALADK 234
              ::||: || |||: :|  :   | : :|| || : |      | : ::|:|| :|| | |::|
Db        199 VDAFAEQLGKLGNALSNTR-LSGLASKLNNLPDLSLAGPGFDAVSGILSVVSASFILSNK 257

Qy        235 NASTGKKVAAGFELSNQVIGNVTKAISSYVLAQRVAAGLSTTGAVAALITSSIMLAISPL 294
              :|  | | ||| |:| :::||: ||:| |::||||||||||  |     || | : ||||||
Db        258 DADAGTKAAAGIEIESTKILGNIGKAVSQYIIAQRVAAGLSTTAATGGLIGSVVALAISPL 317

Qy        295 AFMNAADKFNHANALDEFAKQFRKFGYDGDHLLAEYQRGVGTIEASLTTISTALGAVSAG 354
              :|:| ||||  |   |::::::|:||||:|| |||  :  |  |||:||||:: | | |||
Db        318 SFLNVADKFERAKQLEQYSERFKKFGYEGDSLLASFYRETGAIEAALTTINSVLSAASAG 377

Qy        355 VSAAAVGSAVGTPIALLVAGVTGLISGILEASKQAMFESVANRLQGKILEWEKQNGGQNY 414
              | ||| || || |:| ||: :||:|||||:||||:|| || :|  || ||||:: |:||
Db        378 VGAAATGSLVGAPVAALVSAITGIISGILDASKQAIFERVATKLANKIDEWEKKH-GKNY 436

Qy        415 FDKGYDSRYAAYLANNLKFLSELNKELEAERVIAITQQRWDNNIGELAGITKLGERIKSG 474
              |: |||:|::|:|  :  : ||: |||   |||:||||||| ||||||||: |   |||
Db        437 FENGYDARHSAFLEDTFELLSQYNKEYSVERVVAITQQRWDVNIGELAGITRKGSDTKSG 496

Qy        475 KAYADAFEDGKKVEAG----SNITLDAKTGIIDISNSNGKKTQALHFTSPLLTAGTESRE 530
              ||| |  ||:|| :|                      |   | ||:|: |  ||  |  | :|: ||| | ||
Db        497 KAYVDFFEEGKLLEKEPDRFDKKVFDPLEGKIDLSSIN--KTTLLKFVTPVFTAGEEIRE 554

Qy        531 RLTNGKYSYINKLKFGRVKNWQVTDGEASSKL-DFSKVIQRVAETEGTDEIGLIVNAKAG 589
              |   ||| |: :|      : | ||  :: : : |:: :||     : :| ::  : : :   |
Db        555 RKQTGKYEYMTELFVKGKEKWVVTGVQSHNAIYDYTNLIQLAIDKKG-EKRQVTIESHLG 613

Qy        590 --NDDIFVGQGKMNIDGGDGHDRVFYSK-DGGFGNITVDGTSATEAGSYTVNRKV-ARGD 645
                  || |::   |    :   |:|||   :| |  |  | |:  :| ||  || :|| |  | |::: |
Db        614 EKNDRIYLSSGSSIVYAGNGHDVAYYDKTDTGY--LTFDGQSAQKAGEYIVTKELKADVK 671

Qy        646 IYHEVVKRQETKVGKRTETIQYRDYELR--KVGYGYQSTDNLKSVEEVIGSQFNDVFKGS 703
              :  ||||  |:   ||||:| ::||||||    ::|  | ::  |  |  ||||:|||   |  | |
Db        672 VLKEVVKTQDISVGKRSEKLEYRDYELSPFELGNGIRAKDELHSVEEIIGSNRKDKFFGS 731

Qy        704 KFNDIFHSGEGD-----------------DLLDGGAGDDRLFGGKGNDRLSGDEGDDLL 745
              :| ||||  :||                 |:: || |:| | || ||||| | :|:: |
Db        732 RFTDIFHGAKGDDEIYGNDGHDILYGDDGNDVIHGGDGNDHLVGGNGNDRLIGGKGNNFL 791

Qy        746 DGGSGDD----------VLNGGAGNDVY-------IFRKGDGNDTLYDGTGND------- 781
               :|| |||            || ||||||:            :|  | | ||| :| | | |
Db        792 NGGDGDDELQVFEGQYNVLLGGAGNDILYGSDGTNLFDGGVGNDKIYGGLGKDIYRYSKE 851

Qy        782 ------------------KLAFADANI----SDIMIERTKEGIIVKRNDHSG-SINIPR 817
                                 |:| |         |::: :      | :       |::| :: |
Db        852 YGRHIIIEKGGDDDTLLLSDLSFKDVGFIRIGDDLLVNKRIGGTLYYHEDYNGNALTIKD 911

Qy        818 WYITSNLQNYQSNKTDHKIEQLIGKDGSYITSDQIDKILQDKKDGTVITSQELKKLADEN 877
              |:    :  |:|   |||:::  |||:|: |  : ::       :            || |     |
Db        912 WF--KEGKEGQNN----KIEKIVDKDGAYVLSQYLTELTAPGRGINYFNGLEEKLYYGEG 965

Qy        878 KSQKLSASDIASSLNKLVGSMALF-GTANSVSSNALQPI 915
                :   :   :  : :::  |     |  |    ||  :   |:
Db        966 YN---ALPQLRKDIEQIISSTGAFTGDHGKVSVGSGGPL 1001


RESULT 36
AAY51406
ID   AAY51406 standard; protein; 1022 AA.
XX
AC   AAY51406;
XX
DT   05-MAY-2000  (first entry)
XX
DE   A. pleuropneumoniae apxIA protein.
XX
KW   RTX toxin; vaccine; Repeats in Toxins; immunostimulatory; antimicrobial;
KW   antiinflammatory; antiarthritic; antiabortive; treatment; pneumonia;
KW   pleuropneumonia; septicemia; nephritis; arthritis; endocarditis;
KW   shipping fever; abortion; whooping cough; sleepy foal disease; joint ill;
KW   urinary infection; peritonitis; meningitis; gastroenteritis;
KW   passive immunization; apxIA.
XX
OS   Actinobacillus pleuropneumoniae.
XX
PN   US6019984-A.
XX
PD   01-FEB-2000.
XX
PF   23-DEC-1996;    96US-0772270.
XX
```

```
PR   01-MAR-1995;   95US-0396244.
XX
PA   (UYGU-) UNIV GUELPH.
XX
PI   Mallard B,  Rosendal S,  MacInnes J,  Ricciatti P;
XX
DR   WPI; 2000-146864/13.
DR   N-PSDB; AAZ88584.
XX
PT   Bacterial preparation comprising microorganisms which produce a member
PT   of the Repeats in Toxins (RTX) family, useful for treating swine
PT   pleuropneumonia, arthritis in swine, shipping fever and abortion in
PT   cattle, and sleepy foal disease -
XX
PS   Disclosure; Column 47-52; 96pp; English.
XX
CC   This invention describes a novel bacterial preparation (I) which
CC   comprises one or more isolated and purified strain(s) of a microorganism,
CC   cultured in tryptone yeast extract (TYE) broth, which produces one or
CC   more RTX toxins (belonging to the family of toxins referred to as Repeats
CC   in Toxins), where the strain(s) have at least one RTX toxin which is
CC   cell-associated. The products of the invention have immunostimulatory,
CC   antimicrobial, antiinflammatory, antiarthritic and antiabortive activity.
CC   The bacterial preparation may be used as vaccines for the prophylaxis and
CC   treatment of infectious diseases caused by strains of microorganisms
CC   which produce one or more RTX toxins. The infectious diseases are swine
CC   pleuropneumonia, pneumonia, septicemia, nephritis and arthritis in swine;
CC   septicemia, nephritis, endocarditis and arthritis in piglets; shipping
CC   fever and abortion in cattle; whooping cough, sleepy foal disease or
CC   joint ill (purulent nephritis, arthritis) in foals; septicemia,
CC   polyarthritis and abortion in horses; and urinary infections,
CC   peritonitis, meningitis, and gastroenteritis. The bacterial preparations
CC   may also be used to prepare antibodies which may be used as a means of
CC   passive immunization. This sequence represents the Actinobacillus
CC   pleuropneumoniae apxIA protein described in the method of the invention.
XX
SQ   Sequence   1022 AA;

  Query Match             37.9%;  Score 1763.5;  DB 21;  Length 1022;
  Best Local Similarity   40.0%;  Pred. No. 2.4e-101;
  Matches  400;  Conservative  181;  Mismatches  309;  Indels  109;  Gaps   26;

Qy          8 KSNIQAGLNSTKSGL-------KNLYLAIPKDYDPQKGGTLNDFIKAADELGIARLAEEP 60
              ||  ::|  : |:||        : | | |||||    | :||| :|||: |||     |
Db         21 KSAAKSGAGALKNGLGQVKQAGQKLILYIPKDYQASTGSSLNDLVKAAEALGIEVHRSEK 80

Qy         61 NHTETAKKSVDTVNQFLSLTQTGIAISATKLEKFLQKHSTNKLAKGL-DSVENIDRKLGK 119
              | |  ||:    |  : |   :: |||: | : :| | |:    ||:| |  | | : ::| |
Db         81 NGTALAKELFGTTEKLLGFSERGIALFAPQFDKLLNKN--QKLSKSLGGSSEALGQRLNK 138

Qy        120 ASNVLSTLSSFLGTALAGIELDSLIKK----GDAAPDALAKASIDLINEIIGNLSQSTQT 175
                  || |  ||||||:||::||||:::        | :     |||| :||   ::: |:: :| |
Db        139 TQTALSALQSFLGTAIAGMDLDSLLRRRRNGEDVSGSELAKAGVDLAAQLVDNIASATGT 198

Qy        176 IEAFSSQLAKLGSTISQAKGFSNIGNKLQNL-NFSKTNLGLEIITGLLSGISAGFALADK 234
              ::||: || |||: :|  :   | : :|| || : |     | : ::|:|| :|| | |::|
Db        199 VDAFAEQLGKLGNALSNTR-LSGLASKLNNLPDLSLAGPGFDAVSGILSVVSASFILSNK 257

Qy        235 NASTGKKVAAGFELSNQVIGNVTKAISSYVLAQRVAAGLSTTGAVAALITSSIMLAISPL 294
              :|   |  | ||| |:| :::|: ||:| |::||||||||||  |     || |  : ||||||
Db        258 DADAGTKAAAGIEISTKILGNIGKAVSQYIIAQRVAAGLSTTAATGGLIGSVVALAISPL 317

Qy        295 AFMNAADKFNHANALDEFAKQFRKFGYDGDHLLAEYQRGVGTIEASLTTISTALGAVSAG 354
              :|:| ||||  |    |::::::|:||||:|| |||  :  |   | |||:||||:: |  | |||
Db        318 SFLNVADKFERAKQLEQYSERFKKFGYEGDSLLASFYRETGAIEAALTTINSVLSAASAG 377

Qy        355 VSAAAVGSAVGTPIALLVAGVTGLISGILEASKQAMFESVANRLQGKILEWEKQNGGQNY 414
               | ||| || || ||  |:| ||: :||:|||||:||||:|| || :|  || ||||:: |:||
Db        378 VGAAATGSLVGAPVAALVSAITGIISGILDASKQAIFERVATKLANKIDEWEKKH-GKNY 436

Qy        415 FDKGYDSRYAAYLANNLKFLSELNKELEAERVIAITQQRWDNNIGELAGITKLGERIKSG 474
              |: |||:|::|:|  :   : ||: |||   |||:|||||||| ||||||||: |     |||
Db        437 FENGYDARHSAFLEDTFELLSQYNKEYSVERVVAITQQRWDVNIGELAGITRKGSDTKSG 496

Qy        475 KAYADAFEDGKKVEAG----SNITLDAKTGIIDISNSNGKKTQALHFTSPLLTAGTESRE 530
              ||| | ||:|| :|              |    | ||:|: |  ||  |  | :|: ||| | ||
Db        497 KAYVDFFEEGKLLEKEPDRFDKKVFDPLEGKIDLSSIN--KTTLLKFVTPVFTAGEEIRE 554

Qy        531 RLTNGKYSYINKLKFGRVKNWQVTDGEASSKL-DFSKVIQRVAETEGTDEIGLIVNAKAG 589
                |  ||| |: :|     : | ||  :: : : |:: :||    : :|  ::  : : :  |
Db        555 RKQTGKYEYMTELFVKGKEKWVVTGVQSHNAIYDYTNLIQLAIDKKG-EKRQVTIESHLG 613

Qy        590 --NDDIFVGQGKMNIDGGDGHDRVFYSK-DGGFGNITVDGTSATEAGSYTVNRKV-ARGD 645
                  || |:: |   :  |:|||  :| | | |:  :| || || :|| | | ::: |
Db        614 EKNDRIYLSSGSSIVYAGNGHDVAYYDKTDTGY--LTFDGQSAQKAGEYIVTKELKADVK 671

Qy        646 IYHEVVKRQETKVGKRTETIQYRDYELR--KVGYGYQSTDNLKSVEEVIGSQFNDVFKGS 703
              : ||||  |:   ||||:| ::||||||     ::| | :: | | ||||:|||    | | ||
Db        672 VLKEVVKTQDISVGKRSEKLEYRDYELSPFELGNGIRAKDELHSVEEIIGSNRKDKFFGS 731

Qy        704 KFNDIFHSGEGD-----------------DLLDGGAGDDRLFGGKGNDRLSGDEGDDLL 745
              :| ||||  :||                   |:: || |:| | || ||||| |:|:: |
Db        732 RFTDIFHGAKGDDEIYGNDGHDILYGDDGNDVIHGGDGNDHLVGGNGNDRLIGGKGNNFL 791

Qy        746 DGGSGDD----------VLNGGAGNDVY-------IFRKGDGNDTLYDGTGND------- 781
              :|| |||           || ||||||:       :|  | ||| :| | | |
Db        792 NGGDGDDELQVFEGQYNVLLGGAGNDILYGSDGTNLFDGGVGNDKIYGGLGKDIYRYSKE 851

Qy        782 ------------------KLAFADANI----SDIMIERTKEGIIVKRNDHSG-SINIPR 817
                                |:| |      |::: :   | :     |::| :: |
Db        852 YGRHIIIEKGGDDDTLLLSDLSFKDVGFIRIGDDLLVNKRIGGTLYYHEDYNGNALTIKD 911

Qy        818 WYITSNLQNYQSNKTDHKIEQLIGKDGSYITSDQIDKILQDKKDGTVITSQELKKLADEN 877
              |:    :  |:|    |||:::  |||:|:  |  : ::     :         ||   |
Db        912 WF--KEGKEGQNN----KIEKIVDKDGAYVLSQYLTELTAPGRGINYFNGLEEKLYYGEG 965

Qy        878 KSQKLSASDIASSLNKLVGSMALF-GTANSVSSNALQPI 915
                :   :   : ::: |   |  |     ||  :  |:
Db        966 YN---ALPQLRKDIEQIISSTGAFTGDHGKVSVGSGGPL 1001


RESULT 37
AAR76991
ID   AAR76991 standard; Protein; 1023 AA.
XX
AC   AAR76991;
XX
DT   13-MAR-1996  (first entry)
XX
DE   LhaA (low homology to appA) protein.
XX
KW   RTX; repeat; vaccine; antigenic; immunisation; pneumonia;
KW   Ap1 disease.
XX
OS   Actinobacillus pleuropneumoniae.
XX
FH   Key             Location/Qualifiers
FT   misc_difference 750
FT                   /note = "corresponds to ATT codon"
FT   misc_difference 922
FT                   /note = "corresponds to ATT codon"
XX
PN   JP07138185-A.
XX
PD   30-MAY-1995.
XX
PF   23-JUN-1993;    93JP-0152264.
XX
PR   23-JUN-1993;    93JP-0152264.
XX
PA   (NISK ) NIPPON SEIBUTSU KAGAKU KENKYUSHO ZH.
XX
DR   WPI; 1995-228639/30.
DR   N-PSDB; AAQ94783, AAQ94784.
XX
PT   A vaccine contg. a product of the LhaA gene as the active ingredient
```

```
PT   - for prophylaxis of Actinobacillus (Haemophilus) pleuropneumoniae
PT   infectious diseases.
XX
PS   Claim 3; Page 8-11; 15pp; Japanese.
XX
CC   AAR76991 represents the lhaA (low homology to appA) gene product.
CC   The lhaA protein (which is a toxin component and contains a repeat
CC   region) and fragments of this protein are useful as the active
CC   ingredient in vaccines for the prophylaxis of Actinobacillus
CC   (Haemophilus) pleuropneumoniae infectious (Apl) diseases.
XX
SQ   Sequence   1023 AA;

  Query Match             37.2%;  Score 1727;  DB 16;  Length 1023;
  Best Local Similarity   41.4%;  Pred. No. 4.6e-99;
  Matches  390;  Conservative  169;  Mismatches  271;  Indels  112;  Gaps   26;

Qy          8 KSNIQAGLNSTKSGL-------KNLYLAIPKDYDPQKGGTLNDFIKAADELGIARLAEEP 60
              || ::| : |:||        : | | ||||    | :||| :|||: |||    |
Db         21 KSAAKSGAGALKNGLGQVKQAGQKLILYIPKDYQASTGSSLNDLVKAAEALGIEVHRSEK 80

Qy         61 NHTETAKKSVDTVNQFLSLTQTGIAISATKLEKFLQKHSTNKLAKGL-DSVENIDRKLGK 119
              | | ||:   | : |  :: |||: | : :| | |:    ||:| |  | | : ::| |
Db         81 NGTALAKELFGTTEKLLGFSERGIALFAPQFDKLLNKN--QKLSKSLGGSSEALGQRLNK 138

Qy        120 ASNVLSTLSSFLGTALAGIELDSLIKK----GDAAPDALAKASIDLINEIIGNLSQSTQT 175
                  || | ||||||:||::||||:::      | :    |||| :||   ::: |:: :| |
Db        139 TQTALSALQSFLGTAIAGMDLDSLLRRRRNGEDVSGSELAKAGVDLAAQLVDNIASATGT 198

Qy        176 IEAFSSQLAKLGS---TISQAKGFSNIGNKLQNL-NFSKTNLGLEIITGLLSGISAGFAL 231
              ::||: || ||    |::     | : :|| || : |    | : ::|:|| :|| | |
Db        199 VDAFAEQLGKLAMPYLTLA----LSGLASKLNNLPDLSLAGPGFDAVSGILSVVSASFIL 254

Qy        232 ADKNASTGKKVAAGFELSNQVIGNVTKAISSYVLAQRVAAGLSTTGAVAALITSSIMLAI 291
              ::|:| | | ||| |:| :::||: ||:| |::||||||||||| |    || | : |||
Db        255 SNKDADAGTKAAAGIEISTKILGNIGKAVSQYIIAQRVAAGLSTTAATGGLIGSVVALAI 314

Qy        292 SPLAFMNAADKFNHANALDEFAKQFRKFGYDGDHLLAEYQRGVGTIEASLTTISTALGAV 351
              |||:|:| ||||   | |:::::|:||||:|| ||| : |    | |||:||||.: | |
Db        315 SPLSFLNVADKFERAKQLEQYSERFKKFGYEGDSLLASFYRETGAIEAALTTINSVLSAA 374

Qy        352 SAGVSAAAVGSAVGTPIALLVAGVTGLISGILEASKQAMFESVANRLQGKILEWEKQNGG 411
              |||| ||| || || |:| ||: :||:|||||:|||||:|| || :|   || ||||:: |
Db        375 SAGVGAAATGSLVGAPVAALVSAITGIISGILDASKQAIFERVATKLANKIDEWEKKH-G 433

Qy        412 QNYFDKGYDSRYAAYLANNLKFLSELNKELEAERVIAITQQRWDNNIGELAGITKLGERI 471
              :|||: |||:|::|:|  :  :  ||: |||   |||:|||||||| |||||||||: |
Db        434 KNYFENGYDARHSAFLEDTFELLSQYNKEYSVERVVAITQQRWDVNIGELAGITRKGADA 493

Qy        472 KSGKAYADAFEDGKKVEAG----SNITLDAKTGIIDISNSNGKKTQALHFTSPLLTAGTE 527
              |||||| ||:|| :|               |  | ||:|:: |  || | | |: ||| |
Db        494 KSGKAYVDFFEEGKLLEKDPDRFDKKVFDPLEGKIDLSSIN--KTTLLKFITPVFTAGEE 551

Qy        528 SRERLTNGKYSYINKLKFGRVKNWQVTDGEASSKL-DFSKVIQRVAETEGTDEIGLIVNA 586
                 |||  ||| |: :|        |  |  :|:: : : :|: :  : :| ::  : : :
Db        552 IRERKQTGKYEYMTELFVKGKEKWVVTGVESHNAIYDYTNLIQLAIDKKG-EKRQVTIES 610

Qy        587 KAG--NDDIFVGQGKMNIDGGDGHDRVFYSK-DGGFGNITVDGTSATEAGSYTVNRKV-A 642
                 |  || |:: |    :   |:|||  :| | | |: :| || || :|| | | ::: |
Db        611 HLGEKNDRIYLSSGSSIVYAGNGHDVAYYDKTDTGY--LTFDGQSAQKAGEYIVTKELKA 668

Qy        643 RGDIYHEVVKRQETKVGKR-TETIQYRDYELR--KVGYGYQSTDNLKSVEEVIGSQFNDV 699
                 :  |||| |:  |||  :| ::||||||    ::| | :: | | ||||:|||   |
Db        669 DVKVLKEVVKTQDISVGKTCSEKLEYRDYELSPFELGNGIRAKDELHSVEEIIGSNRKDK 728

Qy        700 FKGSKFNDIFHSGEGD-----------------DLLDGGAGDDRLFGGKGNDRLSGDEG 741
              | ||:| |||| :||                  |:: || |:| | || ||||| | :|
Db        729 FFGSRFTDIFHGAKGDDEIYGNDGHDILYGDDGNDVIHGGDGNDHLVGGNGNDRLIGGKG 788

Qy        742 DDLLDGGSGDD----------VLNGGAGNDVY-------IFRKGDGNDTLYDGTGND--- 781
              :: |:|| |||          || ||||||:         :| |||  | ||| :| | |
Db        789 NNFLNGGDGDDELQVFEGQYNVLLGGAGNDILYGSDGTNLFDGGVGNDKIYGGLGKDIYR 848

Qy        782 ----------------------KLAFADANI----SDIMIERTKEGIIVKRNDHSG-SI 813
                                    |:| |      |::: :  | :   |::| ::
Db        849 YSKEYGRHIIIEKGGDDDTLLLSDLSFKDVGFIRIGDDLLVNKRIGGTLYYHEDYNGNAL 908

Qy        814 NIPRWYITSNLQNYQSNKTDHKIEQLIGKDGSYITSDQIDKI 855
               | |:   :  |:|   |||::: |||:|: |  : ::
Db        909 TIKDWF--KEGKEGQNN----KIEKIVDKDGAYVLSQYLTEL 944


RESULT 38
AAE04636
ID   AAE04636 standard; Protein; 608 AA.
XX
AC   AAE04636;
XX
DT   10-SEP-2001  (first entry)
XX
DE   Pasteurella haemolytica modified leukotoxin 66 (lkt66) protein.
XX
KW   Leukotoxin 66; lkt66; respiratory disease; infection; therapy;
KW   immunostimulant; antibacterial; vaccine; transgenic plant;
KW   animal feed; mutant; mutein.
XX
OS   Pasteurella haemolytica.
OS   Synthetic.
XX
PN   WO200144289-A2.
XX
PD   21-JUN-2001.
XX
PF   15-DEC-2000; 2000WO-CA01498.
XX
PR   17-DEC-1999;    99US-0172148.
XX
PA   (UYGU-) UNIV GUELPH OFFICE.
XX
PI   Lo RYC,  Shewen PE,  Lee RWH,  Hodgins D,  Strommer JN;
XX
DR   WPI; 2001-408470/43.
DR   N-PSDB; AAD08975.
XX
PT   Modified leukotoxin polypeptide is useful in a vaccine to prevent or
PT   treat Mannheimia (Pasteurella) infection (particularly M. haemolytica
PT   infection), and disease associated with a leukotoxin, e.g., respiratory
PT   disease -
XX
PS   Claim 2; Fig 2; 70pp; English.
XX
CC   The present sequence is Pasteurella (Mannheimia) haemolytica
CC   modified leukotoxin-66 (lkt66) protein. The modification comprises
CC   the removal of amino acids within the hydrophobic transmembrane
CC   domain of a full length leukotoxin protein. Modified leukotoxin
CC   sequences are used in vaccines to treat or prevent diseases associated
CC   with leukotoxin, e.g., respiratory disease, and Mannheimia infection
CC   (particularly M. haemolytica infection). In addition, the vaccine
CC   is used to prepare a medicament. Furthermore, the plant transformed
CC   with modified leukotoxin sequences is fed to an animal such as a
CC   ruminant, to prevent or treat respiratory diseases.
XX
SQ   Sequence   608 AA;

  Query Match             31.5%;  Score 1464.5;  DB 22;  Length 608;
  Best Local Similarity   48.2%;  Pred. No. 5.3e-83;
  Matches  290;  Conservative  105;  Mismatches  180;  Indels   27;  Gaps    9;

Qy        341 LTTISTAL---------GAVSAGVSAAAVGSAVGTPIALLVAGVTGLISGILEASKQAMF 391
              |||:|  |         |   | |     || : :||||||:|:|:|| ||: ||||||
Db          5 LTTLSNGLKNTLTATKSGLHKAGQSLTQAGSVIASPIALLVSGITGVISTILQYSKQAMF 64

Qy        392 ESVANRLQGKILEWEKQNGGQNYFDKGYDSRYAAYLANNLKFLSELNKELEAERVIAITQ 451
               | |||::  ||:||||  | :|||: |||:|||  | :  :|||  ||||:|||||||||
Db         65 EHVANKIHNKIVEWEKNNHGKNYFENGYDARYLANLQDNMKFLLNLNKELQAERVIAITQ 124
```

```
Qy        452 QRWDNNIGELAGITKLGERIKSGKAYADAFEDGKKVEAGSNITLDAKTGIIDISNSNGKK 511
Db        125 QQWDNNIGDLAGISRLGEKVLSGKAYVDAFEEGKHIKADKLVQLDSANGIIDVSNSGKAK 184

Qy        512 TQALHFTSPLLTAGTESRERLTNGKYSYINKLKFGRVKNWQVTDGEASSKLDFSKVIQRV 571
Db        185 TQHILFRTPLLTPGTEHRERVQTGKYEYITKLNINRVDSWKITDGAASSTFDLTNVVQRI 244

Qy        572 -------AETEGTDEIGLIVNAKAGNDDIFVGQGKMNIDGGDGHDRVFYSKDGGFGNITV 624
Db        245 GIELDNAGNVTKTKETKIIAKLGEGDDNVFVGSGTTEIDGGEGYDRVHYSR-GNYGALTI 303

Qy        625 DGTSATEAGSYTVNRKVARGDIYHEVVKRQETKVGKRTETIQYRDYELRKVGYGYQSTDN 684
Db        304 DATKETEQGSYTVNRFVETGKALHEVTSTHTALVGNREEKIEYR-HSNNQHHAGYYTKDT 362

Qy        685 LKSVEEVIGSQFNDVFKGSKFNDIFHSGEGDDLLDGGAGDDRLFGGKGNDRLSGDEGDDL 744
Db        363 LKAVEEIIGTSHNDIFKGSKFNDAFNGGDGVDTIDGNDGNDRLFGGKGDDILDGGNGDDF 422

Qy        745 LDGGSGDDVLNGGAGNDVYIFRKGDGNDTLYDGTGNDKLAFADANISDIMIERTKEGIIV 804
Db        423 IDGGKGNDLLHGGKGDDIFVHRKGDGNDIITDSDGNDKLSFSDSNLKDLTFEKVKHNLVI 482

Qy        805 KRNDHSGSINIPRWY----ITSNLQNYQSNKTDHKIEQLIGKDGSYITSDQIDKILQDKK 860
Db        483 -TNSKKEKVTIQNWFREADFAKEVPNYKATK-DEKIEEIIGQNGERITSKQVDDLI--AK 538

Qy        861 DGTVITSQELKKLADENKSQKLSASDIASSLNKLVGSMALFGTANSVSSNALQPITQPTQ 920
Db        539 GNGKITQDELSKVVDNYELLKHS-KNVTNSLDKLISSVSAFTSSNDSRNVLVAPTSMLDQ 597

Qy        921 GI 922
Db        598 SL 599
```

RESULT 39
AAR86998
ID   AAR86998 standard; Protein; 758 AA.
XX
AC   AAR86998;
XX
DT   04-JUL-1996 (first entry)
XX
DE   Enterohaemorrhagic E.coli hlyA gene product.
XX
KW   Enterohaemorrhagic Escherichia coli; virulent; EHEC; O157:H7 serotype;
KW   detection; probe; primer; hlyA gene; enterohaemorrhagic colitis;
KW   haemolytic uremic syndrome; mesenteric adenitis.
XX
OS   Escherichia coli (enterohaemorrhagic).
XX
PN   US5475098-A.
XX
PD   12-DEC-1995.
XX
PF   14-JUN-1994;   94US-0258188.
XX
PR   14-JUN-1994;   94US-0258188.
XX
PA   (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
PI   Hall RH, Xu JG;
XX
DR   WPI; 1996-048546/05.
DR   N-PSDB; AAT08098.
XX
PT   Entero:haemorrhagic E. coli (EHEC) nucleic acid sequences - useful
PT   for probe and primer design for sensitive and specific detection of
PT   EHEC
XX
PS   Claim 1; Columns 37-42; 32pp; English.
XX
CC   Enterohaemorrhagic E.coli (EHEC) associated with enterohaemorrhagic
CC   colitis, haemolytic uremic syndrome and mesenteric adenitis have
CC   been found to carry a hlyA gene and a hylB gene, separated by an
CC   intergenic region. The hly genes and the intergenic region are
CC   absent from bacteria not associated with these diseases and so
CC   provide a useful target for detecting EHEC pathogens, esp. O157:H7
CC   serotype E.coli. The present sequence is that of the protein
CC   encoded by the EHEC hlyA gene.
XX
SQ   Sequence   758 AA;

```
  Query Match             29.6%;  Score 1376.5;  DB 17;  Length 758;
  Best Local Similarity   40.5%;  Pred. No. 2.2e-77;
  Matches  304;  Conservative  129;  Mismatches  225;  Indels   93;  Gaps   16;

Qy        229 FALADKNASTGKKVAAGFELSNQVIGNVTKAISSYVLAQRVAAGLSTTGAVAALITSSIM 288
Db          2 FILGNSDAHTGTKAAAGIELTTQVLGNVGKAVSQYILAQRMAQGLSTTAASAGLITSAVM 61

Qy        289 LAISPLAFMNAADKFNHANALDEFAKQFRKFGYDGDHLLAEYQRGVGTIEASLTTISTAL 348
Db         62 LAISPLSFLAAADKFERAKQLESYSERFKKLNYEGDALLAAFHKETGAIDAALTTINTVL 121

Qy        349 GAVSAGVSAAAVGSAVGTPIALLVAGVTGLISGILEASKQAMFESVANRLQGKILEWEKQ 408
Db        122 SSVSAGVSAASSASLIGAPISMLVSALTGTISGILEASKQAMFEHVAEKFAARINEWEKE 181

Qy        409 NGGQNYFDKGYDSRYAAYLANNLKFLSELNKELEAERVIAITQQRWDNNIGELAGITKLG 468
Db        182 H-GKNYFENGYDARHAAFLEDSLSLLADFSRQHAVERAVAITQQHWDEKIGELAGITRNA 240

Qy        469 ERIKSGKAYADAFEDGKKVEAG----SNITLDAKTGIIDISNSNGKKTQALHFTSPLLTA 524
Db        241 DRSQSGKAYINYLENGGLLEAQPKEFTQQVFDPQKGTIDL--STGNVSSVLTFITPTFTP 298

Qy        525 GTESRERLTNGKYSYINKLKFGRVKNWQVTDGEASSK--LDFSKVIQRVAETEGTDEIGL 582
Db        299 GEEVRERKQSGKYEYMTSLIVNGKDTWSV-KGIKNHKGVYDYSKLIQFVEKNNKHYQARI 357

Qy        583 IVNAKAGNDDIFVGQGKMNIDGGDGHDRVFYSKDGGFGNITVDGTSATEAGSYTVNRKVA 642
Db        358 ISELGDKDDVVYSGAGSSEVFAGEGYDTVSYNKT-DVGKLTIDATGASKPGEYIVSKNM- 415

Qy        643 RGD--IYHEVVKRQETKVGKRTETIQYRDYELRKVGYGYQSTDNLKSVEEVIGSQFNDVF 700
Db        416 YGDVKVLQEVVKEQEVSVGKRTEKIQYRDFEFRTGGIPYDVIDNLHSVEELIGGKHDDEF 475

Qy        701 KGSKFNDIFHSG-----------------EGDDLLDGGAGDDRLFGGKGNDRLSGDEGD 742
Db        476 KGGKFNDIFHGADGNDYIEGNYGNDRLYGDDGDDYISGGQGDDQLFGGSGNDKLSGGDGN 535

Qy        743 -----------------------------------DLLDGGSGDDVLNGGAGNDVYIFR 766
Db        536 NYLTGGSGNDELQAHGAYNILSGGTGDDKLYGGGGIDLLDGGEGNDYLNGGFGNDIYVYG 595

Qy        767 KGDGNDTLYD-GTGNDKLAFADANISDIMIERTKEGIIVKR----------NDHSGSIN 814
Db        596 QNYGHHTIADEGGKGDRLHLSDISFDDIAFKRVGNDLIMNKAINGVLSFNESNDVNG-IT 654

Qy        815 IPRWYITSNLQNYQSNKTDHKIEQLIGKDGSYITSDQIDKILQDKKDGTVITSQELKKLA 874
Db        655 FKNWFAKD-----ASGADNHLVEVITDKDGREI---KVDKIPHNNNE----RSGYIKASN 702

Qy        875 DENKSQKLSASDIASSLNKLVGSMALFGTAN 905
Db        703 IASEKNMVNITSVANDINKIISSVSGFDSGD 733
```

RESULT 40
AAE04637
ID   AAE04637 standard; Protein; 450 AA.
XX

```
AC   AAE04637;
XX
DT   10-SEP-2001  (first entry)
XX
DE   Pasteurella haemolytica modified leukotoxin 50 (lkt50) protein.
XX
KW   Leukotoxin 50; lkt50; respiratory disease; infection; therapy;
KW   immunostimulant; antibacterial; vaccine; transgenic plant;
KW   animal feed; mutant; mutein.
XX
OS   Pasteurella haemolytica.
OS   Synthetic.
XX
PN   WO200144289-A2.
XX
PD   21-JUN-2001.
XX
PF   15-DEC-2000; 2000WO-CA01498.
XX
PR   17-DEC-1999;   99US-0172148.
XX
PA   (UYGU-) UNIV GUELPH OFFICE.
XX
PI   Lo RYC,  Shewen PE,  Lee RWH,  Hodgins D,  Strommer JN;
XX
DR   WPI; 2001-408470/43.
DR   N-PSDB; AAD08976.
XX
PT   Modified leukotoxin polypeptide is useful in a vaccine to prevent or
PT   treat Mannheimia (Pasteurella) infection (particularly M. haemolytica
PT   infection), and disease associated with a leukotoxin, e.g., respiratory
PT   disease -
XX
PS   Claim 4; Fig 12; 70pp; English.
XX
CC   The present sequence is Pasteurella (Mannheimia) haemolytica
CC   modified leukotoxin-50 (lkt50) protein. The modification comprises
CC   the removal of amino acids within the hydrophobic transmembrane
CC   domain of a full length leukotoxin protein. Modified leukotoxin
CC   sequences are used in vaccines to treat or prevent diseases associated
CC   with leukotoxin, e.g., respiratory disease, and Mannheimia infection
CC   (particularly M. haemolytica infection). In addition, the vaccine
CC   is used to prepare a medicament. Furthermore, the plant transformed
CC   with modified leukotoxin sequences is fed to an animal such as a
CC   ruminant, to prevent or treat respiratory diseases.
XX
SQ   Sequence   450 AA;

  Query Match            25.0%;  Score 1159.5;  DB 22;  Length 450;
  Best Local Similarity  49.6%;  Pred. No. 3.6e-64;
  Matches  225;  Conservative   76;  Mismatches  136;  Indels   17;  Gaps    7;

Qy        433 FLSELNKELEAERVIAITQQRWDNNIGELAGITKLGERIKSGKAYADAFEDGKKVEAGSN 492
              ||  |||||:||||||||||:||||||:||||::|||::  |||||  ||||:||  ::|
Db          1 FLLNLNKELQAERVIAITQQQWDNNIGDLAGISRLGEKVLSGKAYVDAFEEGKHIKADKL 60

Qy        493 ITLDAKTGIIDISNSNGKKTQALHFTSPLLTAGTESRERLTNGKYSYINKLKFGRVKNWQ 552
              : ||:  ||||:|||   ||| :  | :||||  ||| |||:  |||  || ||   || :|:
Db         61 VQLDSANGIIDVSNSGKAKTQHILFRTPLLTPGTEHRERVQTGKYEYITKLNINRVDSWK 120

Qy        553 VTDGEASSKLDFSKVIQRV-------AETEGTDEIGLIVNAKAGNDDIFVGQGKMNIDGG 605
              :||| |||   | : |:||:                 |  |   :|     |:|::|||  |     ||||
Db        121 ITDGAASSTFDLTNVVQRIGIELDNAGNVTKTKETKIIAKLGEGDDNVFVGSGTTEIDGG 180

Qy        606 DGHDRVFYSKDGGFGNITVDGTSATEAGSYTVNRKVARGDIYHEVVKRQETKVGKRTETI 665
              :|:|||  ||:  |  :|  :|:|  |    ||  |||||||  |    |   |||          ||  |  |  |
Db        181 EGYDRVHYSR-GNYGALTIDATKETEQGSYTVNRFVETGKALHEVTSTHTALVGNREEKI 239

Qy        666 QYRDYELRKVGYGYQSTDNLKSVEEVIGSQFNDVFKGSKFNDIFHSGEGDDLLDGGAGDD 725
              :||  :     :     ||  :  |  ||:|||:||:   ||:||||||||  |:  |:|  |  :||   |:|
Db        240 EYR-HSNNQHHAGYYTKDTLKAVEEIIGTSHNDIFKGSKFNDAFNGGDGVDTIDGNDGND 298

Qy        726 RLFGGKGNDRLSGDEGDDLLDGGSGDDVLNGGAGNDVYIFRKGDGNDTLYDGTGNDKLAF 785
              |||||||:|  |  |   |||  :|||  |:|:|:||  |:|:::  ||||||  :  |   |||||:|
Db        299 RLFGGKGDDILDGGNGDDFIDGGKGNDLLHGGKGDDIFVHRKGDGNDIITDSDGNDKLSF 358

Qy        786 ADANISDIMIERTKEGIIVKRNDHSGSINIPRWY----ITSNLQNYQSNKTDHKIEQLIG 841
              :|:|:  |:   |:   |    :::   |              :  |    |:               :  ||::  |  |  |||::||
Db        359 SDSNLKDLTFEKVKHNLVI-TNSKKEKVTIQNWFREADFAKEVPNYKATK-DEKIEEIIG 416

Qy        842 KDGSYITSDQIDKILQDKKDGTVITSQELKKLAD 875
              ::|   |||  |:|  ::        |          ||   ||  |:  |
Db        417 QNGERITSKQVDDLI--AKGNGKITQDELSKVVD 448

Search completed: February 17, 2004, 10:13:13
Job time : 55 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: February 17, 2004, 10:13:20 ; Search time 47 Seconds
(without alignments)
3130.628 Million cell updates/sec

Title: US-10-069-799-5
Perfect score: 927
Sequence: 1 MSNINVIKSNIQAGLNSTKS..........SSNALQPITQPTQGILAPSV 927

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 1107863 seqs, 158726573 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : A_Geneseq_19Jun03:*
1: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2002.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	927	100.0	927	22	AAB62110	M. bovis Dalton 2d
2	14	1.5	758	17	AAR86998	Enterohaemorrhagic
3	14	1.5	1049	18	AAW22159	ApxIIIB protein.
4	14	1.5	1049	21	AAY51412	A. pleuropneumonia
5	14	1.5	1244	15	AAR54781	Leukotoxin AppIIIA
6	11	1.2	956	12	AAR12561	APPA haemolysin an
7	11	1.2	956	18	AAW22156	ApxIIC protein. A
8	11	1.2	956	21	AAY51410	A. pleuropneumonia
9	11	1.2	1022	18	AAW22152	ApxIA protein. Ac
10	11	1.2	1022	21	AAY51406	A. pleuropneumonia
11	11	1.2	1023	16	AAR76991	LhaA (low homology
12	10	1.1	450	22	AAE04637	Pasteurella haemol
13	10	1.1	490	20	AAY33929	Leukotoxin carrier
14	10	1.1	544	17	AAW03943	LKT-GnRH protein f
15	10	1.1	544	19	AAW79570	LKT-GnRH chimeric
16	10	1.1	608	22	AAE04636	Pasteurella haemol
17	10	1.1	695	19	AAW79573	LKT-GnRH chimeric
18	10	1.1	695	21	AAY58361	Leukotoxin/gonadot
19	10	1.1	695	21	AAY58133	Gonadotropin relea
20	10	1.1	802	22	ABB67818	Drosophila melanog
21	10	1.1	924	12	AAR10889	Leukotoxin 352 enc
22	10	1.1	924	14	AAR42385	Recombinant leukot
23	10	1.1	924	14	AAR42380	Recombinant leukot
24	10	1.1	924	14	AAR42378	Recombinant leukot
25	10	1.1	926	12	AAR14482	LKT352. Pasteurel
26	10	1.1	926	14	AAR34545	Leukotoxin 352 pro
27	10	1.1	926	15	AAR50291	Recombinant leukot
28	10	1.1	926	17	AAW03945	P. haemolytica tru
29	10	1.1	926	19	AAW79568	Leukotoxin 352 pol
30	10	1.1	934	17	AAW07637	P. suis leukotoxin
31	10	1.1	936	14	AAR34547	GnRH-leukotoxin ge
32	10	1.1	943	14	AAR34546	Somatostatin-leuko
33	10	1.1	951	14	AAR34548	Rotavirus VP4-leuk
34	10	1.1	953	11	AAR07167	105kD PTX protein
35	10	1.1	953	12	AAR15159	Leukotoxin from P.
36	10	1.1	953	14	AAR43865	Leukotoxin protein
37	10	1.1	953	15	AAR60072	PtxA protein of Pa
38	10	1.1	953	22	AAE04638	Pasteurella haemol
39	10	1.1	977	17	AAW03942	LKT-GnRH protein f
40	10	1.1	977	19	AAW79569	LKT-GnRH chimeric
41	10	1.1	1069	15	AAR52748	Bovine IFNgamma/LK
42	10	1.1	1069	18	AAW13867	Chimeric protein #
43	10	1.1	1069	21	AAB21074	Bovine gamma-IFN/P
44	10	1.1	1098	13	AAR22103	Bovine IL-2 - LKT
45	10	1.1	1098	15	AAR52747	Bovine IL-2/LKT ch
46	10	1.1	1098	18	AAW13866	Chimeric protein #
47	10	1.1	1098	21	AAB21073	Bovine IL-2/Pasteu
48	10	1.1	1334	12	AAR14481	LktA:lacZ fusion p
49	10	1.1	1334	15	AAR50290	lktA::lacZ fusion
50	10	1.1	1403	12	AAR10890	lktA::lacZ fusion
51	10	1.1	1403	15	AAR54629	Mannuronan C-5-epi
52	10	1.1	1403	19	AAW79574	Leukotoxin 101. S
53	8	0.9	73	22	AAM86900	Human immune/haema
54	8	0.9	97	23	ABP07319	Human ORFX protein
55	8	0.9	186	21	AAY75008	Neisseria gonorrhe
56	8	0.9	190	22	AAU67626	Propionibacterium
57	8	0.9	251	21	AAY75011	Neisseria gonorrhe
58	8	0.9	251	24	ABP77888	N. gonorrhoeae ami
59	8	0.9	262	18	AAW21779	Protein encoded by
60	8	0.9	341	23	ABP66327	Bifidobacterium lo
61	8	0.9	368	23	ABB49818	Listeria monocytog
62	8	0.9	480	23	ABB93674	Herbicidally activ
63	8	0.9	553	15	AAR54628	Mannuronan C-5-epi
64	8	0.9	580	23	ABR38824	A. niger tripeptid
65	8	0.9	766	22	AAG82742	S. epidermidis ope
66	8	0.9	805	23	ABP39976	Staphylococcus epi
67	8	0.9	825	22	ABB59301	Drosophila melanog
68	8	0.9	872	15	AAR54631	Fragment of mannur
69	8	0.9	997	15	AAR54630	Mannuronan C-5-epi
70	8	0.9	1262	18	AAW13505	B. bronchiseptica
71	8	0.9	1644	18	AAW13504	B. bronchiseptica
72	7	0.8	12	12	AAR12283	Parvovirus B19 epi
73	7	0.8	25	15	AAR53953	Peptide reactive w
74	7	0.8	34	23	AAO21905	BASB204 related pe
75	7	0.8	34	23	AAO21906	BASB204 related pe
76	7	0.8	35	23	AAO21919	BASB204 related pe
77	7	0.8	38	23	AAO21901	BASB204 related pe
78	7	0.8	38	23	AAO21903	BASB204 related pe
79	7	0.8	38	23	AAO21907	BASB204 related pe
80	7	0.8	38	23	AAO21908	BASB204 related pe
81	7	0.8	38	23	AAO21909	BASB204 related pe
82	7	0.8	38	23	AAO21910	BASB204 related pe

83	7	0.8	39	23	AAO21920	BASB204 related pe
84	7	0.8	44	21	AAB09392	Hepatitis GB virus
85	7	0.8	49	23	ABP42813	Human ovarian anti
86	7	0.8	59	23	AAU83459	Parvovirus VP1up c
87	7	0.8	61	23	AAO21918	BASB204 related pr
88	7	0.8	62	22	AAU55937	Propionibacterium
89	7	0.8	62	23	AAO21911	BASB204 related pr
90	7	0.8	62	23	ABP03526	Human ORFX protein
91	7	0.8	62	23	ABP04702	Human ORFX protein
92	7	0.8	63	22	AAU55510	Propionibacterium
93	7	0.8	63	23	AAU90973	Transplant media a
94	7	0.8	67	22	AAU21093	Human novel foetal
95	7	0.8	69	21	AAG18350	Arabidopsis thalia
96	7	0.8	76	23	AAO21912	BASB204 related pr
97	7	0.8	80	19	AAW44949	T. serrulatus scor
98	7	0.8	80	23	AAO21915	BASB204 related pr
99	7	0.8	81	21	AAG33666	Arabidopsis thalia
100	7	0.8	84	23	AAO21913	BASB204 related pr
101	7	0.8	87	21	AAG24373	Arabidopsis thalia
102	7	0.8	87	21	AAG45657	Arabidopsis thalia
103	7	0.8	88	23	AAO21914	BASB204 related pr
104	7	0.8	88	24	ABR47812	Human secreted pro
105	7	0.8	89	20	AAY38428	Human secreted pro
106	7	0.8	89	20	AAY38413	Human secreted pro
107	7	0.8	89	23	AAO21916	BASB204 related pr
108	7	0.8	89	23	AAO21917	BASB204 related pr
109	7	0.8	90	23	ABP33780	Human ORF2753 prot
110	7	0.8	91	23	ABU51050	Helicobacter pylor
111	7	0.8	91	23	ABP08943	Human ORFX protein
112	7	0.8	92	21	AAB52604	Helicobacter pylor
113	7	0.8	98	21	AAG24372	Arabidopsis thalia
114	7	0.8	98	21	AAG45656	Arabidopsis thalia
115	7	0.8	103	20	AAY00227	Enterococcus faeca
116	7	0.8	103	22	AAU35515	Haemophilus influe
117	7	0.8	103	23	ABP43446	E faecalis EF117 a
118	7	0.8	103	24	ABU13725	Enterococcus faeca
119	7	0.8	104	21	AAB15964	E. coli proliferat
120	7	0.8	104	22	AAU34755	E. coli cellular p
121	7	0.8	104	22	AAU38514	Salmonella typhi c
122	7	0.8	104	22	AAG98249	Escherichia coli p
123	7	0.8	104	22	AAG98853	E. coli growth and
124	7	0.8	106	23	AAO17370	Human breast susce
125	7	0.8	106	23	ABP08476	Human ORFX protein
126	7	0.8	109	22	AAU48874	Propionibacterium
127	7	0.8	109	23	ABG90944	Human colon specif
128	7	0.8	109	23	ABJ04319	Human colon specif
129	7	0.8	113	23	ABG77231	Selected Interacti
130	7	0.8	114	22	ABG55762	Human liver peptid
131	7	0.8	114	22	ABB40460	Peptide #7966 enco
132	7	0.8	114	22	ABB24806	Protein #6805 enco
133	7	0.8	114	22	AAM61282	Human brain expres
134	7	0.8	114	22	AAM74005	Human bone marrow
135	7	0.8	114	22	AAM20191	Peptide #6625 enco
136	7	0.8	114	22	AAM34180	Peptide #8217 enco
137	7	0.8	114	22	AAB94807	Human protein sequ
138	7	0.8	114	23	ABG43900	Human peptide enco
139	7	0.8	115	21	AAG03210	Human secreted pro
140	7	0.8	115	23	ABP42840	Human ovarian anti
141	7	0.8	118	22	AAM25357	Human protein sequ
142	7	0.8	121	19	AAW80409	A secreted protein
143	7	0.8	121	23	ABP61827	Human polypeptide
144	7	0.8	122	23	AAU81790	Gossypium hirsutum
145	7	0.8	123	21	AAY81131	Neisseria gonorrho
146	7	0.8	123	21	AAY81132	Neisseria meningit
147	7	0.8	123	24	ABP80703	N. gonorrhoeae ami
148	7	0.8	127	23	ABP66092	Bifidobacterium lo
149	7	0.8	128	20	AAY00226	Enterococcus faeca
150	7	0.8	128	22	AAU28029	Novel human secret
151	7	0.8	128	23	ABP43445	E faecalis EF117 p
152	7	0.8	128	24	ABU13724	Enterococcus faeca
153	7	0.8	129	22	AAU61084	Propionibacterium
154	7	0.8	129	24	ABU00043	Human novel polype
155	7	0.8	133	21	AAG24371	Arabidopsis thalia

156	7	0.8	133	21	AAG45655	Arabidopsis thalia
157	7	0.8	134	21	AAB43577	Human cancer assoc
158	7	0.8	134	22	AAM80842	Human haematologic
159	7	0.8	134	22	AAM81621	Human haematologic
160	7	0.8	134	22	AAM81862	Human haematologic
161	7	0.8	135	6	AAP50637	Salmon pleopiomela
162	7	0.8	136	21	AAG01455	Human secreted pro
163	7	0.8	138	18	AAW27248	Pseudomonas fluore
164	7	0.8	140	22	AAU30632	Novel human secret
165	7	0.8	140	22	AAM40427	Human polypeptide
166	7	0.8	141	22	ABB61586	Drosophila melanog
167	7	0.8	141	22	AAB90817	Human shear stress
168	7	0.8	142	24	ABU00295	Human novel polype
169	7	0.8	143	23	ABP51338	Human MDDT SEQ ID
170	7	0.8	143	23	ABG60186	Human DITHP polype
171	7	0.8	146	23	ABP03724	Human ORFX protein
172	7	0.8	148	22	AAB63453	Human breast cance
173	7	0.8	151	21	AAB28200	Human xs99. Homo
174	7	0.8	151	22	AAM78494	Human protein SEQ
175	7	0.8	155	22	AAU31740	Novel human secret
176	7	0.8	159	21	AAB33138	Eucalyptus grandis
177	7	0.8	162	23	ABB97818	Human secretory po
178	7	0.8	163	22	AAB63282	Human breast cance
179	7	0.8	165	23	ABP30875	Streptococcus poly
180	7	0.8	168	22	ABB64189	Drosophila melanog
181	7	0.8	171	23	ABJ11102	Yeast selected int
182	7	0.8	172	22	AAU53524	Propionibacterium
183	7	0.8	174	23	ABB90413	Human polypeptide
184	7	0.8	175	23	ABP25846	Streptococcus poly
185	7	0.8	179	23	AAE24889	Pneumococcus sp. c
186	7	0.8	180	22	AAM79478	Human protein SEQ
187	7	0.8	183	21	AAG54968	Arabidopsis thalia
188	7	0.8	185	22	AAM42209	Human polypeptide
189	7	0.8	189	21	AAB43461	Human cancer assoc
190	7	0.8	190	22	AAU87397	Novel central nerv
191	7	0.8	190	22	AAU87673	Novel central nerv
192	7	0.8	190	22	AAU17297	Novel signal trans
193	7	0.8	194	21	AAG18057	Arabidopsis thalia
194	7	0.8	203	23	ABB89137	Human polypeptide
195	7	0.8	207	22	ABB67746	Drosophila melanog
196	7	0.8	210	12	AAR13401	Parvo virus B19 PC
197	7	0.8	218	22	ABG19282	Novel human diagno
198	7	0.8	219	22	AAU36380	Pseudomonas aerugi
199	7	0.8	219	23	ABU51331	Helicobacter pylor
200	7	0.8	221	21	AAG29207	Arabidopsis thalia
201	7	0.8	226	21	AAG29206	Arabidopsis thalia
202	7	0.8	227	12	AAR13400	Parvo virus B19 PA
203	7	0.8	227	20	AAY23228	Erythrovirus V9 VP
204	7	0.8	227	23	AAU83477	B19 human parvovir
205	7	0.8	235	21	AAG10686	Arabidopsis thalia
206	7	0.8	241	22	AAU36060	Klebsiella pneumon
207	7	0.8	242	21	AAG10303	Arabidopsis thalia
208	7	0.8	243	21	AAG54718	Arabidopsis thalia
209	7	0.8	244	21	AAG18056	Arabidopsis thalia
210	7	0.8	246	22	ABG10652	Novel human diagno
211	7	0.8	250	12	AAR13404	Parvo virus B19 PA
212	7	0.8	253	22	ABG20679	Novel human diagno
213	7	0.8	256	22	AAU35071	Enterococcus faeca
214	7	0.8	257	23	AAE20548	Corn diverged delt
215	7	0.8	258	21	AAG14028	Arabidopsis thalia
216	7	0.8	258	22	ABG27045	Novel human diagno
217	7	0.8	260	22	ABB59673	Drosophila melanog
218	7	0.8	261	22	AAU51721	Propionibacterium
219	7	0.8	269	23	ABB89792	Human polypeptide
220	7	0.8	274	21	AAB32523	S. lavendulae MmcJ
221	7	0.8	275	21	AAB58399	Lung cancer associ
222	7	0.8	275	21	AAG08359	Arabidopsis thalia
223	7	0.8	276	18	AAW20242	H. pylori transpor
224	7	0.8	276	18	AAW24615	H. pylori transpor
225	7	0.8	278	21	AAG08358	Arabidopsis thalia
226	7	0.8	280	23	ABP30416	Streptococcus poly
227	7	0.8	281	23	ABB49721	Listeria monocytog
228	7	0.8	283	23	AAU79170	Human voltage-depe

229	7	0.8	284	24	ABP57913	X. campestris avrX
230	7	0.8	293	22	AAB94712	Human protein sequ
231	7	0.8	293	24	ABU70387	Human adipocyte Se
232	7	0.8	296	22	AAM39415	Human polypeptide
233	7	0.8	296	22	AAB92555	Human protein sequ
234	7	0.8	296	24	ABP60400	Pseudomonas fluore
235	7	0.8	300	21	AAG14027	Arabidopsis thalia
236	7	0.8	301	23	ABP27439	Streptococcus poly
237	7	0.8	302	21	AAG46712	Arabidopsis thalia
238	7	0.8	302	22	AAG82984	S. epidermidis ope
239	7	0.8	305	21	AAG10685	Arabidopsis thalia
240	7	0.8	305	24	AAE34574	Human CD47 protein
241	7	0.8	307	24	ABJ19021	Pathogen specific
242	7	0.8	308	22	AAG92563	C glutamicum prote
243	7	0.8	310	23	ABR38856	A. niger aspergill
244	7	0.8	311	21	AAG29205	Arabidopsis thalia
245	7	0.8	311	22	AAB79756	Corynebacterium gl
246	7	0.8	311	23	ABB77561	Corynebacteria Cys
247	7	0.8	314	21	AAG08357	Arabidopsis thalia
248	7	0.8	315	23	ABP39676	Staphylococcus epi
249	7	0.8	319	23	ABB79408	Mouse neuronal reg
250	7	0.8	321	22	AAB65977	Human secreted pro
251	7	0.8	321	22	AAB64717	Gene 5 human secre
252	7	0.8	323	23	ABP62901	Human polypeptide
253	7	0.8	323	23	AAE17874	Sequence of ompK36
254	7	0.8	324	22	AAU36286	Pseudomonas aerugi
255	7	0.8	325	21	AAG21863	Arabidopsis thalia
256	7	0.8	325	21	AAG47535	Arabidopsis thalia
257	7	0.8	325	23	ABU50952	Helicobacter pylor
258	7	0.8	326	21	AAB52500	Helicobacter pylor
259	7	0.8	328	21	AAG14026	Arabidopsis thalia
260	7	0.8	329	20	AAW86356	Partial human DNAX
261	7	0.8	329	23	AAE16097	Human DNAX Toll li
262	7	0.8	331	22	ABB71976	Drosophila melanog
263	7	0.8	331	22	ABG20249	Novel human diagno
264	7	0.8	331	23	AAU74332	Human cytoskeleton
265	7	0.8	333	23	AAO22146	Ramoplanin biosynt
266	7	0.8	336	22	ABB64015	Drosophila melanog
267	7	0.8	340	21	AAG46711	Arabidopsis thalia
268	7	0.8	341	23	ABB54677	Lactococcus lactis
269	7	0.8	343	20	AAY56769	C. trachomatis ser
270	7	0.8	343	21	AAG46710	Arabidopsis thalia
271	7	0.8	353	22	AAG75406	Human colon cancer
272	7	0.8	354	23	ABU51076	Helicobacter pylor
273	7	0.8	358	23	ABU05789	M. tuberculosis an
274	7	0.8	362	22	AAU23384	Novel human enzyme
275	7	0.8	365	21	AAG10302	Arabidopsis thalia
276	7	0.8	366	23	AAU72942	Neisseria meningit
277	7	0.8	366	24	ABP77317	N. gonorrhoeae ami
278	7	0.8	368	23	ABP26740	Streptococcus poly
279	7	0.8	373	23	ABP41917	Human ovarian anti
280	7	0.8	379	22	AAU18518	Human cytoskeletal
281	7	0.8	379	22	AAU18543	Human cytoskeletal
282	7	0.8	380	20	AAY06627	C. elegans tyrosyl
283	7	0.8	380	21	AAY84308	A tyrosylprotein s
284	7	0.8	380	23	AAE20551	Corn diverged delt
285	7	0.8	381	23	ABB93951	Herbicidally activ
286	7	0.8	382	19	AAW98312	H. pylori GHPO 886
287	7	0.8	382	22	ABB63575	Drosophila melanog
288	7	0.8	382	22	AAM41201	Human polypeptide
289	7	0.8	382	24	ABU58164	Corn stress respon
290	7	0.8	383	21	AAG10301	Arabidopsis thalia
291	7	0.8	386	19	AAW98237	H. pylori GHPO 699
292	7	0.8	386	19	AAW71489	Helicobacter polyp
293	7	0.8	386	22	AAU69494	Human purified sec
294	7	0.8	387	23	ABU51114	Helicobacter pylor
295	7	0.8	392	22	ABB65237	Drosophila melanog
296	7	0.8	394	18	AAW20797	H. pylori transpor
297	7	0.8	396	24	ABP99303	Orthosomycin biosy
298	7	0.8	396	24	ABP99350	Orthosomycin biosy
299	7	0.8	396	24	ABP76716	Streptomyces virid
300	7	0.8	400	21	AAY82775	Human chordin rela
301	7	0.8	400	21	AAY53033	Human secreted pro

302	7	0.8	407	23	ABB54596	Lactococcus lactis
303	7	0.8	409	22	ABG19403	Novel human diagno
304	7	0.8	409	22	AAG82454	S. epidermidis ope
305	7	0.8	409	22	AAG82606	S. epidermidis ope
306	7	0.8	415	23	ABU51693	Helicobacter pylor
307	7	0.8	418	22	ABG21503	Novel human diagno
308	7	0.8	424	22	ABB61353	Drosophila melanog
309	7	0.8	424	23	ABB07924	Fruit fly pellino
310	7	0.8	424	23	ABP40164	Staphylococcus epi
311	7	0.8	429	22	ABG20995	Novel human diagno
312	7	0.8	435	23	ABJ11075	Yeast selected int
313	7	0.8	435	23	ABG77204	Selected Interacti
314	7	0.8	439	22	AAB04085	Botulism toxin C f
315	7	0.8	439	23	ABU50795	Helicobacter pylor
316	7	0.8	440	21	AAY77135	Synthetic botulinu
317	7	0.8	440	22	AAB04091	Botulism toxin hea
318	7	0.8	440	23	ABJ11271	Yeast selected int
319	7	0.8	440	23	ABG77369	Selected Interacti
320	7	0.8	448	23	ABB54422	Lactococcus lactis
321	7	0.8	451	22	AAU01689	Gene 31 human secr
322	7	0.8	454	22	ABB52906	Escherichia coli p
323	7	0.8	455	22	AAU35965	Helicobacter pylor
324	7	0.8	458	22	AAU35728	Helicobacter pylor
325	7	0.8	458	22	AAB96223	Putative P. abyssi
326	7	0.8	460	22	ABG17038	Novel human diagno
327	7	0.8	460	22	AAE05398	Human interleukin-
328	7	0.8	460	22	AAE05401	Human IRAK-4 mutan
329	7	0.8	460	22	AAE05403	Human IRAK-4 polym
330	7	0.8	460	22	AAE05404	Human IRAK-4 polym
331	7	0.8	460	23	AAE24859	Human IRAK4 protei
332	7	0.8	460	23	AAE24865	Human IRAK4 protei
333	7	0.8	463	20	AAY36773	Chlamydia trachoma
334	7	0.8	468	22	AAU71895	C. glutamicum meta
335	7	0.8	468	22	AAB79667	Corynebacterium gl
336	7	0.8	468	22	AAB76671	Corynebacterium gl
337	7	0.8	471	18	AAW16794	Aspartase-470. Es
338	7	0.8	472	19	AAW68394	Clostridium botuli
339	7	0.8	477	21	AAY68317	Escherichia coli B
340	7	0.8	478	20	AAY49079	Escherichia coli a
341	7	0.8	478	20	AAY33962	E. coli aspartase
342	7	0.8	478	20	AAY05731	E. coli native asp
343	7	0.8	478	21	AAG47163	Arabidopsis thalia
344	7	0.8	478	21	AAY79297	Escherichia coli a
345	7	0.8	478	23	ABU50702	Helicobacter pylor
346	7	0.8	478	23	ABB09551	Escherichia coli K
347	7	0.8	479	22	ABG10527	Novel human diagno
348	7	0.8	487	22	AAB82711	Aranicola proteoly
349	7	0.8	492	24	AAO16701	Escherichia coli a
350	7	0.8	499	21	AAB01363	1-deoxy-D-xylulose
351	7	0.8	500	21	AAG21939	Arabidopsis thalia
352	7	0.8	500	24	AAE34903	Mouse calcium acti
353	7	0.8	503	8	AAP70219	Sequence of serrap
354	7	0.8	504	22	AAU14800	Novel bone marrow
355	7	0.8	512	22	ABB57797	Drosophila melanog
356	7	0.8	519	23	AAE25289	Human nucleic acid
357	7	0.8	519	23	ABG66662	Human fat responsi
358	7	0.8	520	23	ABG66661	Rat fat responsive
359	7	0.8	521	19	AAW37934	Mouse cAMP regulat
360	7	0.8	523	22	AAB65800	RPP5-like protein
361	7	0.8	551	22	AAG90677	C glutamicum prote
362	7	0.8	551	22	AAB76764	Corynebacterium gl
363	7	0.8	557	19	AAW71586	Human hepatocyte n
364	7	0.8	558	24	ABP77082	N. gonorrhoeae ami
365	7	0.8	560	15	AAR62523	Hookworm anticoagu
366	7	0.8	575	21	AAG21938	Arabidopsis thalia
367	7	0.8	578	21	AAB52575	Helicobacter pylor
368	7	0.8	582	21	AAG47162	Arabidopsis thalia
369	7	0.8	585	22	AAU35539	Haemophilus influe
370	7	0.8	587	22	ABB52856	Escherichia coli p
371	7	0.8	587	24	ABB82486	M. sexta serotonin
372	7	0.8	589	22	ABU52814	Human metabolism-a
373	7	0.8	592	22	AAM93825	Human polypeptide,
374	7	0.8	594	24	ABU06042	N. meningitidis va

375	7	0.8	598	21	AAG21937	Arabidopsis thalia
376	7	0.8	609	19	AAW71581	Human hepatocyte n
377	7	0.8	617	18	AAW27247	Pseudomonas fluore
378	7	0.8	619	23	ABU50976	Helicobacter pylor
379	7	0.8	619	23	ABU52060	Helicobacter pylor
380	7	0.8	621	22	ABB67394	Drosophila melanog
381	7	0.8	622	22	ABB60638	Drosophila melanog
382	7	0.8	627	22	AAM38995	Human polypeptide
383	7	0.8	630	22	ABB63386	Drosophila melanog
384	7	0.8	630	22	ABB67395	Drosophila melanog
385	7	0.8	630	22	ABB67396	Drosophila melanog
386	7	0.8	631	22	ABB66645	Drosophila melanog
387	7	0.8	631	23	AAU12044	Clostridium diffic
388	7	0.8	632	23	ABB49748	Listeria monocytog
389	7	0.8	634	22	ABB65156	Drosophila melanog
390	7	0.8	636	23	ABB07503	Human GTP-binding
391	7	0.8	639	22	ABG25372	Novel human diagno
392	7	0.8	641	23	AAM49040	Human testicular d
393	7	0.8	641	24	AAO16116	Human cancer-testi
394	7	0.8	647	22	ABB60758	Drosophila melanog
395	7	0.8	654	23	ABB54057	Lactococcus lactis
396	7	0.8	660	23	ABB94293	Chlamydia trachoma
397	7	0.8	661	22	ABG30261	Novel human diagno
398	7	0.8	670	23	ABP26115	Streptococcus poly
399	7	0.8	671	21	AAB07763	Amino acid sequenc
400	7	0.8	672	21	AAB21233	Corn MFP1. Zea ma
401	7	0.8	674	22	AAU14364	Human novel protei
402	7	0.8	674	22	AAB95048	Human protein sequ
403	7	0.8	674	23	ABB76698	Human eukaryotic i
404	7	0.8	687	24	ABB98844	Human NEPHA-SV1.PR
405	7	0.8	687	24	ABB98851	Human Eph receptor
406	7	0.8	689	23	ABB49714	Listeria monocytog
407	7	0.8	710	22	AAU33871	Staphylococcus aur
408	7	0.8	711	19	AAW44842	Staphylococcus aur
409	7	0.8	711	22	AAU36597	Staphylococcus aur
410	7	0.8	732	22	AAB61140	Human NOV10 protei
411	7	0.8	737	24	ABB98847	Human NEPHA-SV4.PR
412	7	0.8	737	24	ABB98854	Human Eph receptor
413	7	0.8	754	21	AAB35987	Sorbitol dehydroge
414	7	0.8	772	19	AAW29666	Homo sapiens DH130
415	7	0.8	776	22	ABG28762	Novel human diagno
416	7	0.8	781	16	AAW08986	Human parvovirus V
417	7	0.8	781	20	AAY23227	Erythrovirus V9 VP
418	7	0.8	781	24	ABP57263	Human parvovirus B
419	7	0.8	781	24	ABP57266	Human parvovirus B
420	7	0.8	784	22	ABB63844	Drosophila melanog
421	7	0.8	812	22	AAG78604	Lawsonia intracell
422	7	0.8	814	22	ABB68374	Drosophila melanog
423	7	0.8	817	14	AAR34942	Human epidermal tr
424	7	0.8	817	23	AAM49673	Human TGK protein.
425	7	0.8	831	24	ABP78978	N. gonorrhoeae ami
426	7	0.8	833	22	ABB68116	Drosophila melanog
427	7	0.8	836	22	AAB62026	Recombinant P. fur
428	7	0.8	839	22	ABB63119	Drosophila melanog
429	7	0.8	843	22	ABB69333	Drosophila melanog
430	7	0.8	848	22	AAB04082	Botulism toxin hea
431	7	0.8	855	22	AAB96787	Putative P. abyssi
432	7	0.8	860	22	ABG20680	Novel human diagno
433	7	0.8	865	22	ABG25368	Novel human diagno
434	7	0.8	866	22	ABB71210	Drosophila melanog
435	7	0.8	868	22	ABG30260	Novel human diagno
436	7	0.8	872	22	ABB70960	Drosophila melanog
437	7	0.8	873	23	AAE25018	Human drug metabol
438	7	0.8	875	24	ABP96075	Human protein kina
439	7	0.8	877	24	ABP96076	Human protein kina
440	7	0.8	888	22	ABB61813	Drosophila melanog
441	7	0.8	892	22	AAB62179	Human p100 protein
442	7	0.8	893	22	AAG67396	Amino acid sequenc
443	7	0.8	893	22	AAB98890	Novel human (NHP)
444	7	0.8	893	23	AAU76324	Human protein kina
445	7	0.8	898	21	AAY83939	Benzene ring hydro
446	7	0.8	899	22	ABB62277	Drosophila melanog
447	7	0.8	929	22	ABB70857	Drosophila melanog
448	7	0.8	969	22	ABB58321	Drosophila melanog
449	7	0.8	974	23	AAU76874	Human EphA full le
450	7	0.8	982	22	ABB71714	Drosophila melanog
451	7	0.8	986	22	ABB57810	Drosophila melanog
452	7	0.8	1000	22	AAM47209	Human NOV3 protein
453	7	0.8	1001	23	AAG68224	Nilaparvata lugens
454	7	0.8	1006	22	ABB60312	Drosophila melanog
455	7	0.8	1008	24	ABB98843	Human NEPHA. Homo
456	7	0.8	1009	22	AAU03553	Human protein kina
457	7	0.8	1009	23	AAE19158	Human kinase polyp
458	7	0.8	1014	20	AAY38743	Neisseria meningit
459	7	0.8	1015	20	AAY38744	Neisseria meningit
460	7	0.8	1022	23	AAU76875	Human EphA extende
461	7	0.8	1030	23	ABP35638	Fungal ZBC protein
462	7	0.8	1032	23	AAE16107	Human DNAX Toll li
463	7	0.8	1034	22	AAU34738	E. coli cellular p
464	7	0.8	1036	21	AAY82776	Human chordin rela
465	7	0.8	1036	21	AAY53034	Human secreted pro
466	7	0.8	1036	22	AAU07141	Human CRIM1 protei
467	7	0.8	1036	22	AAU12242	Human PRO4330 poly
468	7	0.8	1036	23	AAE18852	Human pharmaceutic
469	7	0.8	1036	24	ABU66640	Human PRO polypept
470	7	0.8	1036	24	ABU66916	Human secreted/tra
471	7	0.8	1036	24	ABU59721	Novel secreted and
472	7	0.8	1040	22	ABB23027	Protein #5026 enco
473	7	0.8	1040	22	AAM31141	Peptide #5178 enco
474	7	0.8	1040	23	ABG40604	Human peptide enco
475	7	0.8	1041	20	AAY41768	Human PRO286 prote
476	7	0.8	1041	20	AAY05867	Human Toll protein
477	7	0.8	1041	21	AAB44324	Human PRO286 prote
478	7	0.8	1041	23	AAO21584	Murine Toll-like r
479	7	0.8	1041	23	AAO21586	Murine Toll-like r
480	7	0.8	1041	24	ABU61154	Human PRO286 polyp
481	7	0.8	1059	23	AAO21587	Murine Toll-like r
482	7	0.8	1070	21	AAY93308	A manganese supero
483	7	0.8	1070	22	ABG12878	Novel human diagno
484	7	0.8	1075	19	AAW64472	Human secreted pro
485	7	0.8	1075	22	AAB90748	Human DL162_2 pro
486	7	0.8	1083	20	AAY25169	Human RSC ligase p
487	7	0.8	1095	21	AAY93311	A manganese supero
488	7	0.8	1100	22	AAG84930	Shrimp white spot
489	7	0.8	1109	20	AAY02376	Polypeptide identi
490	7	0.8	1138	11	AAR06461	BtPGS1245 protoxin
491	7	0.8	1138	14	AAR37213	B.t. toxin HD511.
492	7	0.8	1138	14	AAR37214	B.t. toxin HD867.
493	7	0.8	1138	15	AAR46225	Bacillus thuringie
494	7	0.8	1138	15	AAR46226	Bacillus thuringie
495	7	0.8	1160	22	ABG28234	Novel human diagno
496	7	0.8	1247	22	AAG90997	C glutamicum prote
497	7	0.8	1275	24	ABJ25467	Aspergillus fumiga
498	7	0.8	1287	16	AAR79944	Helicobacter pylor
499	7	0.8	1288	18	AAW55547	H. pylori ORF 14ee
500	7	0.8	1288	18	AAW55685	H. pylori ORF 07ee
501	7	0.8	1290	19	AAW98269	H. pylori GHPO 374
502	7	0.8	1291	19	AAW68392	Clostridium botuli
503	7	0.8	1296	14	AAR41198	CT. Helicobacter
504	7	0.8	1400	22	AAM78576	Human protein SEQ
505	7	0.8	1400	22	AAM79560	Human protein SEQ
506	7	0.8	1435	24	ABJ26067	Aspergillus fumiga
507	7	0.8	1476	22	ABB58706	Drosophila melanog
508	7	0.8	1497	22	ABG20797	Novel human diagno
509	7	0.8	1600	22	AAG84976	Shrimp white spot
510	7	0.8	1605	22	ABB70375	Drosophila melanog
511	7	0.8	1841	21	AAB07562	Protein encoded by
512	7	0.8	1908	24	ABJ25817	Aspergillus fumiga
513	7	0.8	2022	24	ABJ26417	Aspergillus fumiga
514	7	0.8	2099	24	ABJ25632	Aspergillus fumiga
515	7	0.8	2162	21	AAB07560	Protein encoded by
516	7	0.8	2405	24	ABJ26232	Aspergillus fumiga
517	7	0.8	2517	22	ABB65918	Drosophila melanog
518	7	0.8	2519	22	ABG16636	Novel human diagno
519	7	0.8	3010	16	AAR95022	Hepatitis GB virus
520	7	0.8	3135	15	AAR57474	P. falciparum tran

521	7	0.8	3135	21	AAB18223	Plasmodium falcipa
522	7	0.8	3234	23	ABP70132	Human NOV39a. Hom
523	7	0.8	3433	24	ABP70647	Amino acid sequenc
524	7	0.8	3433	24	ABB98821	West Nile virus st
525	7	0.8	4999	23	AAO22158	Ramoplanin biosynt
526	7	0.8	19938	24	ABB98398	Streptomyces virid
527	6	0.6	8	22	AAB78491	SIV and HIV gp 41
528	6	0.6	9	22	AAU24331	Human MHC class I
529	6	0.6	9	22	AAU24443	Human MHC molecule
530	6	0.6	9	23	AAE26784	Fibrin binding pep
531	6	0.6	10	20	AAY46794	Immunogenic peptid
532	6	0.6	10	20	AAY07312	Collagen assembly
533	6	0.6	10	22	AAU23878	Human MHC class I
534	6	0.6	10	22	AAU23882	Human MHC class I
535	6	0.6	10	22	AAU24213	Human MHC molecule
536	6	0.6	10	22	AAG95769	Human complementar
537	6	0.6	10	22	AAG84417	Arabidopsis thalia
538	6	0.6	10	22	AAG86218	Saccharomyces cere
539	6	0.6	10	22	AAG86454	Saccharomyces cere
540	6	0.6	10	23	ABG98131	Anti-neovascular p
541	6	0.6	10	24	ABU03336	Human expressed pr
542	6	0.6	10	24	ABP74549	Human PSMA epitope
543	6	0.6	12	15	AAR60404	Antiproliferative
544	6	0.6	12	22	AAU28915	MART-1 antigen pep
545	6	0.6	13	15	AAR43970	N-heparan sulphate
546	6	0.6	13	23	ABB79454	Alanine-stabilised
547	6	0.6	13	23	ABB79457	BDC2.5 peptide. U
548	6	0.6	14	22	ABB56594	Human SNP related
549	6	0.6	15	20	AAY07417	HSV-1 TK amino aci
550	6	0.6	15	20	AAY07309	Collagen assembly
551	6	0.6	15	20	AAW97149	HSV-1 TK amino aci
552	6	0.6	15	21	AAB15883	Human chemokine de
553	6	0.6	15	21	AAB14218	Mutant HIV cleavag
554	6	0.6	15	23	ABP59092	Cell division cycl
555	6	0.6	15	23	AAG79657	Development regula
556	6	0.6	15	23	AAE26768	Fibrin binding pep
557	6	0.6	15	23	ABB79475	Human Orc4Lp 47.3
558	6	0.6	15	23	ABG65786	Plant potential PE
559	6	0.6	15	23	AAU77490	N-terminal peptide
560	6	0.6	15	24	ABG99208	Thymidine kinase (
561	6	0.6	16	18	AAW12672	Orange pectin meth
562	6	0.6	16	22	AAB64741	Human secreted pro
563	6	0.6	16	22	AAB11442	Interleukin 12 der
564	6	0.6	17	21	AAB28076	Human secreted pro
565	6	0.6	17	23	AAE23446	Mouse and rat STIM
566	6	0.6	17	23	AAE23447	Mouse STIM2 peptid
567	6	0.6	18	11	AAR06540	Human B-cell simul
568	6	0.6	18	16	AAR70727	Signal peptide use
569	6	0.6	18	16	AAR76687	Signal peptide use
570	6	0.6	18	23	AAE28328	Human DinP protein
571	6	0.6	18	23	ABG61452	Tick Ixolaris seco
572	6	0.6	18	24	ABJ38351	Preferred TALL-1 m
573	6	0.6	18	24	ABP54217	Human lung specifi
574	6	0.6	19	16	AAR86351	HIV-1 protease inh
575	6	0.6	19	16	AAR86352	HIV-1 protease inh
576	6	0.6	19	19	AAW23945	Human herpesvirus
577	6	0.6	19	23	ABG61451	Tick Ixolaris seco
578	6	0.6	20	12	AAR10909	Immunogenic epitop
579	6	0.6	20	12	AAR14484	LKT neutralising e
580	6	0.6	20	13	AAR24125	Leukotoxin epitope
581	6	0.6	20	15	AAR45183	LKT synthetic epit
582	6	0.6	20	18	AAW13868	Synthetic leukotox
583	6	0.6	20	19	AAW67508	IGF-1/IGFBP inhibi
584	6	0.6	20	21	AAB21076	Pasteurella haemol
585	6	0.6	20	21	AAY84052	Amino acid sequenc
586	6	0.6	20	22	AAG63745	Synthetic peptide
587	6	0.6	20	23	AAE28335	Human DinP protein
588	6	0.6	20	23	ABG31560	Interleukin-12 pep
589	6	0.6	20	23	ABG61450	Tick Ixolaris seco
590	6	0.6	20	23	ABB57663	IGFBP-3 binding pe
591	6	0.6	21	16	AAR76686	Signal peptide use
592	6	0.6	21	22	AAB23787	Neisseria gonorrhe
593	6	0.6	21	22	AAG65528	Antimicrobial pept
594	6	0.6	21	22	AAG65544	Antimicrobial pept
595	6	0.6	21	23	ABG61449	Tick Ixolaris seco
596	6	0.6	22	23	ABG61448	Tick Ixolaris seco
597	6	0.6	22	23	AAU69693	Cell death protect
598	6	0.6	23	21	AAY69795	Human interleukin-
599	6	0.6	23	23	ABG61447	Tick Ixolaris seco
600	6	0.6	23	23	ABB53201	Human ORF107 prote
601	6	0.6	24	13	AAR27088	OmpF signal peptid
602	6	0.6	24	15	AAR50167	OmpF leader sequen
603	6	0.6	24	18	AAW11712	E.coli leader sequ
604	6	0.6	24	22	AAE03613	Human AZAD protein
605	6	0.6	24	22	AAB82365	Protein sequence S
606	6	0.6	24	23	ABB76468	Synthetic leader p
607	6	0.6	24	23	ABB76469	Synthetic leader p
608	6	0.6	24	23	ABB76472	Synthetic leader p
609	6	0.6	24	23	ABG61446	Tick Ixolaris seco
610	6	0.6	24	23	AAU69692	Cell death protect
611	6	0.6	25	21	AAY80001	Optimised IgE-CH3
612	6	0.6	25	23	ABG61445	Tick Ixolaris seco
613	6	0.6	26	12	AAR12451	HIBP1-9(307-331) H
614	6	0.6	26	14	AAR40061	Hib OMP P1 peptide
615	6	0.6	26	15	AAR54955	AntoI grass pollen
616	6	0.6	26	21	AAY84094	Synthetic amino ac
617	6	0.6	26	21	AAY86459	Human gene 47-enco
618	6	0.6	26	22	ABG57550	Human liver peptid
619	6	0.6	26	22	ABB42120	Peptide #9626 enco
620	6	0.6	26	22	AAM63004	Human brain expres
621	6	0.6	26	22	AAM75815	Human bone marrow
622	6	0.6	26	22	AAM88302	Human immune/haema
623	6	0.6	26	22	AAM35923	Peptide #9960 enco
624	6	0.6	26	23	ABG61444	Tick Ixolaris seco
625	6	0.6	27	23	ABG61443	Tick Ixolaris seco
626	6	0.6	28	18	AAW12665	Orange pectin meth
627	6	0.6	28	21	AAY84088	Amino acid sequenc
628	6	0.6	28	22	ABG49076	Human liver peptid
629	6	0.6	28	22	ABG58481	Human liver peptid
630	6	0.6	28	22	ABB29077	Peptide #1728 enco
631	6	0.6	28	22	ABB34237	Peptide #1743 enco
632	6	0.6	28	22	ABB43081	Peptide #10587 enc
633	6	0.6	28	22	ABB19673	Protein #1672 enco
634	6	0.6	28	22	ABB26235	Protein #8234 enco
635	6	0.6	28	22	AAM55032	Human brain expres
636	6	0.6	28	22	AAM63980	Human brain expres
637	6	0.6	28	22	AAM67419	Human bone marrow
638	6	0.6	28	22	AAM76800	Human bone marrow
639	6	0.6	28	22	AAM15247	Peptide #1681 enco
640	6	0.6	28	22	AAM27708	Peptide #1745 enco
641	6	0.6	28	22	AAM36906	Peptide #10943 enc
642	6	0.6	28	22	AAM02991	Peptide #1673 enco
643	6	0.6	28	23	ABG37042	Human peptide enco
644	6	0.6	28	23	ABG45958	Human peptide enco
645	6	0.6	28	23	ABG61442	Tick Ixolaris seco
646	6	0.6	29	23	ABG61441	Tick Ixolaris seco
647	6	0.6	30	4	AAP30193	Sequence encoded b
648	6	0.6	30	16	AAR75651	Human placenta der
649	6	0.6	30	17	AAR92064	Recombinant alpha2
650	6	0.6	30	21	AAB39084	Human secreted pro
651	6	0.6	30	21	AAG55396	Arabidopsis thalia
652	6	0.6	30	22	ABG15956	Novel human diagno
653	6	0.6	30	22	AAU00948	Human B7-H3 polype
654	6	0.6	30	22	AAB87416	Human secreted pro
655	6	0.6	30	23	ABG61440	Tick Ixolaris seco
656	6	0.6	31	22	ABG50089	Human liver peptid
657	6	0.6	31	22	ABB30040	Peptide #2691 enco
658	6	0.6	31	22	ABB35211	Peptide #2717 enco
659	6	0.6	31	22	ABB20652	Protein #2651 enco
660	6	0.6	31	22	AAM56043	Human brain expres
661	6	0.6	31	22	AAM68415	Human bone marrow
662	6	0.6	31	22	AAM16230	Peptide #2664 enco
663	6	0.6	31	22	AAM28720	Peptide #2757 enco
664	6	0.6	31	22	AAM03960	Peptide #2642 enco
665	6	0.6	31	23	ABG37988	Human peptide enco
666	6	0.6	31	23	ABG61439	Tick Ixolaris seco

667	6	0.6	32	23	ABG61438	Tick Ixolaris seco
668	6	0.6	32	23	ABP28129	Streptococcus poly
669	6	0.6	33	21	AAY50193	Streptococcus pyog
670	6	0.6	33	22	ABG49558	Human liver peptid
671	6	0.6	33	22	ABB29550	Peptide #2201 enco
672	6	0.6	33	22	ABB34734	Peptide #2240 enco
673	6	0.6	33	22	ABB20146	Protein #2145 enco
674	6	0.6	33	22	AAM55532	Human brain expres
675	6	0.6	33	22	AAM67916	Human bone marrow
676	6	0.6	33	22	AAM15733	Peptide #2167 enco
677	6	0.6	33	22	AAM28238	Peptide #2275 enco
678	6	0.6	33	22	AAM03466	Peptide #2148 enco
679	6	0.6	33	22	AAB48668	Human CD80 constan
680	6	0.6	33	23	ABG37452	Human peptide enco
681	6	0.6	33	23	ABG61437	Tick Ixolaris seco
682	6	0.6	34	14	AAR36607	IL-6 receptor bind
683	6	0.6	34	23	ABG61436	Tick Ixolaris seco
684	6	0.6	35	21	AAY65241	Human 5' EST relat
685	6	0.6	35	23	ABG70821	Mouse myocardin as
686	6	0.6	35	23	ABG61435	Tick Ixolaris seco
687	6	0.6	35	24	ABG74189	Mouse myocardin as
688	6	0.6	36	21	AAB18727	Amino acid sequenc
689	6	0.6	36	22	ABB40753	Peptide #8259 enco
690	6	0.6	36	22	AAM61613	Human brain expres
691	6	0.6	36	22	AAM74405	Human bone marrow
692	6	0.6	36	23	ABJ11196	Yeast selected int
693	6	0.6	36	23	ABG77294	Selected Interacti
694	6	0.6	36	23	ABG61434	Tick Ixolaris seco
695	6	0.6	36	24	ABP79856	N. gonorrhoeae ami
696	6	0.6	37	22	ABG53245	Human liver peptid
697	6	0.6	37	22	ABB38423	Peptide #5929 enco
698	6	0.6	37	22	AAM59033	Human brain expres
699	6	0.6	37	22	AAM71563	Human bone marrow
700	6	0.6	37	22	AAM31861	Peptide #5898 enco
701	6	0.6	37	23	ABG41374	Human peptide enco
702	6	0.6	37	23	ABG61433	Tick Ixolaris seco
703	6	0.6	37	24	ABR47947	Human secreted pro
704	6	0.6	37	24	ABR00179	Human gene 169 enc
705	6	0.6	38	20	AAY01388	Secreted protein e
706	6	0.6	38	21	AAB56549	Human prostate can
707	6	0.6	38	23	AAE28324	Human DinP protein
708	6	0.6	38	23	ABG61432	Tick Ixolaris seco
709	6	0.6	38	23	ABP30565	Streptococcus poly
710	6	0.6	38	24	ABP96039	Mus musculus DGAT
711	6	0.6	39	21	AAY65068	Human 5' EST relat
712	6	0.6	39	22	AAU20686	Human novel foetal
713	6	0.6	39	23	ABG61384	Tick Ixolaris seco
714	6	0.6	39	23	ABG61431	Tick Ixolaris seco
715	6	0.6	39	24	ABP96500	Peptide-linker-bet
716	6	0.6	40	22	AAB99008	Spinach 23K transf
717	6	0.6	40	23	ABG61383	Tick Ixolaris seco
718	6	0.6	40	23	ABG61430	Tick Ixolaris seco
719	6	0.6	41	19	AAW42440	S. pombe MFm1 mati
720	6	0.6	41	20	AAY33240	Human beta-catenin
721	6	0.6	41	23	ABG61382	Tick Ixolaris seco
722	6	0.6	41	23	ABG61429	Tick Ixolaris seco
723	6	0.6	42	23	ABG61381	Tick Ixolaris seco
724	6	0.6	42	23	ABG61428	Tick Ixolaris seco
725	6	0.6	43	20	AAY14018	YhdD repeat sequen
726	6	0.6	43	22	ABG53412	Human liver peptid
727	6	0.6	43	22	ABG00780	Novel human diagno
728	6	0.6	43	22	ABG11360	Novel human diagno
729	6	0.6	43	22	ABB38568	Peptide #6074 enco
730	6	0.6	43	22	AAM59192	Human brain expres
731	6	0.6	43	22	AAM71727	Human bone marrow
732	6	0.6	43	22	AAM32019	Peptide #6056 enco
733	6	0.6	43	23	AAE28333	Human DinP protein
734	6	0.6	43	23	ABG41541	Human peptide enco
735	6	0.6	43	23	ABG61380	Tick Ixolaris seco
736	6	0.6	43	23	ABG61427	Tick Ixolaris seco
737	6	0.6	44	21	AAB24500	Human secreted pro
738	6	0.6	44	22	AAM90069	Human immune/haema
739	6	0.6	44	23	ABG61379	Tick Ixolaris seco

740	6	0.6	44	23	ABG61426	Tick Ixolaris seco
741	6	0.6	44	24	ABP79654	N. gonorrhoeae ami
742	6	0.6	45	21	AAY80018	IgE immunogenic pe
743	6	0.6	45	22	ABG54757	Human liver peptid
744	6	0.6	45	22	ABB39689	Peptide #7195 enco
745	6	0.6	45	22	ABB24351	Protein #6350 enco
746	6	0.6	45	22	AAM60407	Human brain expres
747	6	0.6	45	22	AAM70731	Human bone marrow
748	6	0.6	45	22	AAM73043	Human bone marrow
749	6	0.6	45	22	AAM19816	Peptide #6250 enco
750	6	0.6	45	22	AAM33265	Peptide #7302 enco
751	6	0.6	45	22	AAB66328	D radiodurans glut
752	6	0.6	45	23	ABG40467	Human peptide enco
753	6	0.6	45	23	ABG42887	Human peptide enco
754	6	0.6	45	23	ABG61378	Tick Ixolaris seco
755	6	0.6	45	23	ABG61425	Tick Ixolaris seco
756	6	0.6	46	21	AAB06102	Caenorhabditis ele
757	6	0.6	46	21	AAY80017	IgE immunogenic pe
758	6	0.6	46	22	ABG49456	Human liver peptid
759	6	0.6	46	22	ABG15316	Novel human diagno
760	6	0.6	46	22	ABB29456	Peptide #2107 enco
761	6	0.6	46	22	ABB34633	Peptide #2139 enco
762	6	0.6	46	22	ABB20043	Protein #2042 enco
763	6	0.6	46	22	AAM55423	Human brain expres
764	6	0.6	46	22	AAM67818	Human bone marrow
765	6	0.6	46	22	AAM15628	Peptide #2062 enco
766	6	0.6	46	22	AAM28126	Peptide #2163 enco
767	6	0.6	46	22	AAM03371	Peptide #2053 enco
768	6	0.6	46	23	ABG37361	Human peptide enco
769	6	0.6	46	23	ABG61377	Tick Ixolaris seco
770	6	0.6	46	23	ABG61424	Tick Ixolaris seco
771	6	0.6	47	21	AAG46731	Arabidopsis thalia
772	6	0.6	47	23	ABG61376	Tick Ixolaris seco
773	6	0.6	47	23	ABG61423	Tick Ixolaris seco
774	6	0.6	48	14	AAR34858	Peptide showing cA
775	6	0.6	48	20	AAW88643	Secreted protein e
776	6	0.6	48	22	ABB50410	Human secreted pro
777	6	0.6	48	22	AAB99015	Wheat 33K signal p
778	6	0.6	48	23	ABG61375	Tick Ixolaris seco
779	6	0.6	48	23	ABG61422	Tick Ixolaris seco
780	6	0.6	49	22	AAB87386	Human gene 45 enco
781	6	0.6	49	23	ABG65353	Human albumin fusi
782	6	0.6	49	23	ABG61374	Tick Ixolaris seco
783	6	0.6	49	23	ABG61421	Tick Ixolaris seco
784	6	0.6	50	20	AAY33423	L. brevis ADH pept
785	6	0.6	50	22	AAO21254	Interleukin-2 homo
786	6	0.6	50	22	AAU62599	Propionibacterium
787	6	0.6	50	22	AAO11706	Human polypeptide
788	6	0.6	50	22	AAG77522	Human colon cancer
789	6	0.6	50	23	ABG61373	Tick Ixolaris seco
790	6	0.6	50	23	ABG61420	Tick Ixolaris seco
791	6	0.6	51	21	AAG34384	Arabidopsis thalia
792	6	0.6	51	23	ABG61372	Tick Ixolaris seco
793	6	0.6	51	23	ABP11189	Human ORFX protein
794	6	0.6	52	22	ABG49785	Human liver peptid
795	6	0.6	52	22	AAU58374	Propionibacterium
796	6	0.6	52	22	AAM55757	Human brain expres
797	6	0.6	52	22	AAM68139	Human bone marrow
798	6	0.6	53	18	AAW55593	H. pylori ORF 09cp
799	6	0.6	53	20	AAY12475	Human 5' EST secre
800	6	0.6	53	21	AAB28752	Sequence homologou
801	6	0.6	53	22	AAO21245	Novel interleukin-
802	6	0.6	53	22	AAO21246	Interleukin-2 homo
803	6	0.6	53	22	AAO11297	Human polypeptide
804	6	0.6	53	22	AAG74943	Human colon cancer
805	6	0.6	53	23	ABG61190	Tick Ixolaris amin
806	6	0.6	53	24	ABP77646	N. gonorrhoeae ami
807	6	0.6	54	22	AAO21247	Interleukin-2 homo
808	6	0.6	54	22	AAO21248	Interleukin-2 homo
809	6	0.6	54	22	AAO21249	Interleukin-2 homo
810	6	0.6	54	22	AAO21251	Interleukin-2 homo
811	6	0.6	54	22	AAO21252	Interleukin-2 homo
812	6	0.6	54	22	AAU44002	Propionibacterium

813	6	0.6	54	22	AAU54145	Propionibacterium
814	6	0.6	54	22	AAM86435	Human immune/haema
815	6	0.6	54	23	ABP62961	Human polypeptide
816	6	0.6	54	23	ABG61191	Tick Ixolaris amin
817	6	0.6	54	23	ABP34785	Human ORF3758 prot
818	6	0.6	54	23	ABP01675	Human ORFX protein
819	6	0.6	54	23	ABP04933	Human ORFX protein
820	6	0.6	54	23	ABP06534	Human ORFX protein
821	6	0.6	55	22	AAU45099	Propionibacterium
822	6	0.6	55	22	AAU18585	Human lung antigen
823	6	0.6	55	22	AAU17855	Novel human respir
824	6	0.6	55	23	ABG61192	Tick Ixolaris amin
825	6	0.6	56	22	AAG78035	Piscirickettsia sa
826	6	0.6	56	23	ABG61193	Tick Ixolaris amin
827	6	0.6	57	22	ABB12332	Human bone marrow
828	6	0.6	57	23	ABG61194	Tick Ixolaris amin
829	6	0.6	58	22	AAU41360	Propionibacterium
830	6	0.6	58	22	AAO08269	Human polypeptide
831	6	0.6	58	23	ABG61195	Tick Ixolaris amin
832	6	0.6	58	23	ABP03070	Human ORFX protein
833	6	0.6	58	23	ABP03723	Human ORFX protein
834	6	0.6	59	21	AAB03675	Microfibril associ
835	6	0.6	59	21	AAY76330	Fragment of human
836	6	0.6	59	22	AAU60996	Propionibacterium
837	6	0.6	59	22	ABG26902	Novel human diagno
838	6	0.6	59	22	AAO11340	Human polypeptide
839	6	0.6	59	23	ABP65003	Human protein SEQ
840	6	0.6	59	23	ABG61196	Tick Ixolaris amin
841	6	0.6	60	19	AAW62252	Subtilase HSFURI f
842	6	0.6	60	21	AAB03674	Microfibrillar ass
843	6	0.6	60	21	AAG01532	Human secreted pro
844	6	0.6	60	22	AAU51783	Propionibacterium
845	6	0.6	60	22	AAU64475	Propionibacterium
846	6	0.6	60	23	ABG61197	Tick Ixolaris amin
847	6	0.6	60	23	ABB57406	Human secreted pro
848	6	0.6	61	22	AAU48693	Propionibacterium
849	6	0.6	61	22	AAU22386	Human cardiovascul
850	6	0.6	61	22	AAM90453	Human immune/haema
851	6	0.6	61	23	ABU51140	Helicobacter pylor
852	6	0.6	61	23	ABG61198	Tick Ixolaris amin
853	6	0.6	61	23	ABP39427	Staphylococcus epi
854	6	0.6	61	24	ABU00458	S. pneumoniae type
855	6	0.6	62	21	AAB16636	Bacteriophage 192
856	6	0.6	62	21	AAG55685	Arabidopsis thalia
857	6	0.6	62	21	AAG59382	Arabidopsis thalia
858	6	0.6	62	22	ABB95751	Human testicular a
859	6	0.6	62	22	AAM95045	Human reproductive
860	6	0.6	62	23	ABG61199	Tick Ixolaris amin
861	6	0.6	62	23	ABP33541	Human ORF2514 prot
862	6	0.6	62	23	ABP10510	Human ORFX protein
863	6	0.6	62	23	ABP28116	Streptococcus poly
864	6	0.6	63	17	AAR95434	Interleukin-2 BETA
865	6	0.6	63	20	AAY41340	Human secreted pro
866	6	0.6	63	21	AAG59385	Arabidopsis thalia
867	6	0.6	63	22	AAG89874	C glutamicum prote
868	6	0.6	63	23	ABG61200	Tick Ixolaris amin
869	6	0.6	63	23	ABP08649	Human ORFX protein
870	6	0.6	63	23	ABP10061	Human ORFX protein
871	6	0.6	64	20	AAW88786	Polypeptide fragme
872	6	0.6	64	21	AAG19329	Arabidopsis thalia
873	6	0.6	64	22	AAU48932	Propionibacterium
874	6	0.6	64	22	AAU51830	Propionibacterium
875	6	0.6	64	22	ABB50622	Human secreted pro
876	6	0.6	64	22	AAM82482	Human immune/haema
877	6	0.6	64	22	AAM88282	Human immune/haema
878	6	0.6	64	23	AAE28312	Human DinP protein
879	6	0.6	64	23	ABG61201	Tick Ixolaris amin
880	6	0.6	65	6	AAP50167	Sequence of the si
881	6	0.6	65	20	AAY60268	Human endometrium
882	6	0.6	65	21	AAB52182	Human secreted pro
883	6	0.6	65	21	AAG55684	Arabidopsis thalia
884	6	0.6	65	21	AAG59381	Arabidopsis thalia
885	6	0.6	65	22	AAM86070	Human immune/haema
886	6	0.6	65	23	AAO22521	Small acid-soluble
887	6	0.6	65	23	ABG61202	Tick Ixolaris amin
888	6	0.6	66	22	AAU56053	Propionibacterium
889	6	0.6	66	22	ABG16373	Novel human diagno
890	6	0.6	66	23	AAE28316	Human DinP protein
891	6	0.6	66	23	ABG61203	Tick Ixolaris amin
892	6	0.6	66	23	ABP33629	Human dehydrogenas
893	6	0.6	67	23	ABG61204	Tick Ixolaris amin
894	6	0.6	67	23	ABP31120	Human ORF93 protei
895	6	0.6	67	23	ABP09599	Human ORFX protein
896	6	0.6	68	22	ABB17375	Human nervous syst
897	6	0.6	68	23	ABG61205	Tick Ixolaris amin
898	6	0.6	68	23	ABP25929	Streptococcus poly
899	6	0.6	69	22	ABG56480	Human liver peptid
900	6	0.6	69	22	ABB41037	Peptide #8543 enco
901	6	0.6	69	22	AAM61894	Human brain expres
902	6	0.6	69	22	AAM74696	Human bone marrow
903	6	0.6	69	22	AAM34812	Peptide #8849 enco
904	6	0.6	69	22	AAG91815	C glutamicum prote
905	6	0.6	69	23	AAE28306	Human DinP protein
906	6	0.6	69	23	ABG44507	Human peptide enco
907	6	0.6	69	23	ABG61206	Tick Ixolaris amin
908	6	0.6	70	21	AAG02211	Human secreted pro
909	6	0.6	70	21	AAG03204	Human secreted pro
910	6	0.6	70	22	AAU40286	Propionibacterium
911	6	0.6	70	22	AAU54423	Propionibacterium
912	6	0.6	70	22	ABG23797	Novel human diagno
913	6	0.6	70	22	AAM94363	Human reproductive
914	6	0.6	70	22	AAM24031	Murine EST encoded
915	6	0.6	70	22	AAM42254	Human breast or ov
916	6	0.6	70	23	ABG61207	Tick Ixolaris amin
917	6	0.6	71	20	AAY01550	Antigenic epitope
918	6	0.6	71	20	AAY12033	Human 5' EST secre
919	6	0.6	71	23	ABG61208	Tick Ixolaris amin
920	6	0.6	71	24	ABG76380	Translation domain
921	6	0.6	72	21	AAB33181	Eucalyptus grandis
922	6	0.6	72	21	AAB38336	Human secreted pro
923	6	0.6	72	22	AAU54405	Propionibacterium
924	6	0.6	72	22	AAM94315	Human reproductive
925	6	0.6	72	22	AAB80421	Gene #6 associated
926	6	0.6	72	23	ABG61209	Tick Ixolaris amin
927	6	0.6	73	20	AAY11552	Human 5' EST secre
928	6	0.6	73	22	AAU60392	Propionibacterium
929	6	0.6	73	22	ABG26128	Novel human diagno
930	6	0.6	73	22	ABB15166	Human nervous syst
931	6	0.6	73	23	ABG61210	Tick Ixolaris amin
932	6	0.6	74	20	AAY31399	HEV peptide z12-or
933	6	0.6	74	20	AAY31398	HEV peptide z12-or
934	6	0.6	74	22	AAU45204	Propionibacterium
935	6	0.6	74	22	AAU51521	Propionibacterium
936	6	0.6	74	22	AAE04257	Human gene 4 encod
937	6	0.6	74	23	ABG61211	Tick Ixolaris amin
938	6	0.6	74	23	ABP34424	Human ORF3397 prot
939	6	0.6	74	23	ABP03480	Human ORFX protein
940	6	0.6	74	24	ABU58181	Corn stress respon
941	6	0.6	75	19	AAW62746	Streptococcus pneu
942	6	0.6	75	21	AAG34537	Arabidopsis thalia
943	6	0.6	75	21	AAY55672	M. pneumoniae ycfB
944	6	0.6	75	22	ABG58275	Human liver peptid
945	6	0.6	75	22	AAU46118	Propionibacterium
946	6	0.6	75	22	AAM63762	Human brain expres
947	6	0.6	75	22	AAM36683	Peptide #10720 enc
948	6	0.6	75	23	ABG61212	Tick Ixolaris amin
949	6	0.6	76	22	ABG56080	Human liver peptid
950	6	0.6	76	22	ABB67244	Drosophila melanog
951	6	0.6	76	22	AAU67291	Propionibacterium
952	6	0.6	76	22	ABG08931	Novel human diagno
953	6	0.6	76	22	ABB40632	Peptide #8138 enco
954	6	0.6	76	22	AAU30207	Novel human secret
955	6	0.6	76	22	AAM61491	Human brain expres
956	6	0.6	76	22	AAM74281	Human bone marrow
957	6	0.6	76	22	AAM34393	Peptide #8430 enco
958	6	0.6	76	23	ABB08031	HMM derived consen

959	6	0.6	76	23	ABG61213	Tick Ixolaris amin
960	6	0.6	77	19	AAW77751	Staphylococcus aur
961	6	0.6	77	21	AAG56300	Zea mays protein f
962	6	0.6	77	21	AAG01348	Human secreted pro
963	6	0.6	77	22	AAM96471	Human reproductive
964	6	0.6	77	23	ABB08026	HMM derived consen
965	6	0.6	77	23	ABG61214	Tick Ixolaris amin
966	6	0.6	78	22	AAU22026	Human cardiovascul
967	6	0.6	78	23	ABG61215	Tick Ixolaris amin
968	6	0.6	78	23	ABB49601	Listeria monocytog
969	6	0.6	79	20	AAY74049	Human prostate tum
970	6	0.6	79	21	AAB34973	Gene 46 human secr
971	6	0.6	79	21	AAG13442	Arabidopsis thalia
972	6	0.6	79	22	ABB67938	Drosophila melanog
973	6	0.6	79	22	ABB70482	Drosophila melanog
974	6	0.6	79	22	ABG05325	Novel human diagno
975	6	0.6	79	22	AAM92933	Human digestive sy
976	6	0.6	79	22	AAG91567	C glutamicum prote
977	6	0.6	79	22	AAU20029	Human liver associ
978	6	0.6	79	23	ABU52000	Helicobacter pylor
979	6	0.6	79	23	ABG61216	Tick Ixolaris amin
980	6	0.6	79	23	ABP40890	Human liver antige
981	6	0.6	79	23	ABB06851	Human nGPCR-Seq106
982	6	0.6	79	23	ABP03550	Human ORFX protein
983	6	0.6	80	18	AAW55227	H. pylori ORF 03gp
984	6	0.6	80	20	AAY37446	Amino acid sequenc
985	6	0.6	80	22	AAU44861	Propionibacterium
986	6	0.6	80	23	ABP54797	Human CD8. Homo s
987	6	0.6	80	23	ABG61217	Tick Ixolaris amin
988	6	0.6	80	23	ABP03882	Human ORFX protein
989	6	0.6	80	23	ABP07264	Human ORFX protein
990	6	0.6	81	17	AAR97308	Parietaria allerge
991	6	0.6	81	20	AAY35749	Chlamydia pneumoni
992	6	0.6	81	21	AAG19328	Arabidopsis thalia
993	6	0.6	81	21	AAG34383	Arabidopsis thalia
994	6	0.6	81	22	AAU54299	Propionibacterium
995	6	0.6	81	22	AAM84803	Human immune/haema
996	6	0.6	81	22	AAO13375	Human polypeptide
997	6	0.6	81	23	ABG61218	Tick Ixolaris amin
998	6	0.6	81	24	ABR48471	Human Beferin. Ho
999	6	0.6	82	19	AAW38476	Streptococcus pneu
1000	6	0.6	82	21	AAG19349	Arabidopsis thalia

ALIGNMENTS

```
RESULT 1
AAB62110
ID   AAB62110 standard; Protein; 927 AA.
XX
AC   AAB62110;
XX
DT   29-MAY-2001  (first entry)
XX
DE   M. bovis Dalton 2d RTX toxin A subunit.
XX
KW   Moraxella; antigen; immune response; infection; RTX toxin; vaccine;
KW   antibacterial; A subunit.
XX
OS   Moraxella bovis.
XX
PN   WO200116172-A1.
XX
PD   08-MAR-2001.
XX
PF   31-AUG-2000; 2000WO-AU01048.
XX
PR   31-AUG-1999;    99AU-0002571.
XX
PA   (CSIR ) COMMONWEALTH SCI & IND RES ORG.
PA   (UYME ) UNIV MELBOURNE.
XX
PI   Farn J,  Strugnell R,  Tennent J;
XX
DR   WPI; 2001-235092/24.
DR   N-PSDB; AAF57290.
XX
PT   Novel Moraxella bovis antigen useful in compositions for raising immune
PT   response in an animal, has protease, lipase or hemolysin activity  -
XX
PS   Claim 26; Fig 5; 60pp; English.
XX
CC   The invention relates to new Moraxella bovis antigens and nucleic acid
CC   sequences encoding these antigenic polypeptides. The antigenic
CC   polypeptides and polynucleotides are useful for raising an immune
CC   response in an animal directed against Moraxella, preferably against
CC   M. bovis or M. catarrhalis, and for treating Moraxella infections. The
CC   present sequence represents the amino acid sequence of the A subunit of
CC   the RTX toxin from M. bovis Dalton 2d.
XX
SQ   Sequence   927 AA;

  Query Match             100.0%;  Score 927;  DB 22;  Length 927;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  927;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MSNINVIKSNIQAGLNSTKSGLKNLYLAIPKDYDPQKGGTLNDFIKAADELGIARLAEEP 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MSNINVIKSNIQAGLNSTKSGLKNLYLAIPKDYDPQKGGTLNDFIKAADELGIARLAEEP 60

Qy         61 NHTETAKKSVDTVNQFLSLTQTGIAISATKLEKFLQKHSTNKLAKGLDSVENIDRKLGKA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 NHTETAKKSVDTVNQFLSLTQTGIAISATKLEKFLQKHSTNKLAKGLDSVENIDRKLGKA 120

Qy        121 SNVLSTLSSFLGTALAGIELDSLIKKGDAAPDALAKASIDLINEIIGNLSQSTQTIEAFS 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 SNVLSTLSSFLGTALAGIELDSLIKKGDAAPDALAKASIDLINEIIGNLSQSTQTIEAFS 180

Qy        181 SQLAKLGSTISQAKGFSNIGNKLQNLNFSKTNLGLEIITGLLSGISAGFALADKNASTGK 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 SQLAKLGSTISQAKGFSNIGNKLQNLNFSKTNLGLEIITGLLSGISAGFALADKNASTGK 240

Qy        241 KVAAGFELSNQVIGNVTKAISSYVLAQRVAAGLSTTGAVAALITSSIMLAISPLAFMNAA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 KVAAGFELSNQVIGNVTKAISSYVLAQRVAAGLSTTGAVAALITSSIMLAISPLAFMNAA 300

Qy        301 DKFNHANALDEFAKQFRKFGYDGDHLLAEYQRGVGTIEASLTTISTALGAVSAGVSAAAV 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 DKFNHANALDEFAKQFRKFGYDGDHLLAEYQRGVGTIEASLTTISTALGAVSAGVSAAAV 360

Qy        361 GSAVGTPIALLVAGVTGLISGILEASKQAMFESVANRLQGKILEWEKQNGGQNYFDKGYD 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GSAVGTPIALLVAGVTGLISGILEASKQAMFESVANRLQGKILEWEKQNGGQNYFDKGYD 420

Qy        421 SRYAAYLANNLKFLSELNKELEAERVIAITQQRWDNNIGELAGITKLGERIKSGKAYADA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 SRYAAYLANNLKFLSELNKELEAERVIAITQQRWDNNIGELAGITKLGERIKSGKAYADA 480

Qy        481 FEDGKKVEAGSNITLDAKTGIIDISNSNGKKTQALHFTSPLLTAGTESRERLTNGKYSYI 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 FEDGKKVEAGSNITLDAKTGIIDISNSNGKKTQALHFTSPLLTAGTESRERLTNGKYSYI 540

Qy        541 NKLKFGRVKNWQVTDGEASSKLDFSKVIQRVAETEGTDEIGLIVNAKAGNDDIFVGQGKM 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 NKLKFGRVKNWQVTDGEASSKLDFSKVIQRVAETEGTDEIGLIVNAKAGNDDIFVGQGKM 600

Qy        601 NIDGGDGHDRVFYSKDGGFGNITVDGTSATEAGSYTVNRKVARGDIYHEVVKRQETKVGK 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 NIDGGDGHDRVFYSKDGGFGNITVDGTSATEAGSYTVNRKVARGDIYHEVVKRQETKVGK 660

Qy        661 RTETIQYRDYELRKVGYGYQSTDNLKSVEEVIGSQFNDVFKGSKFNDIFHSGEGDDLLDG 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 RTETIQYRDYELRKVGYGYQSTDNLKSVEEVIGSQFNDVFKGSKFNDIFHSGEGDDLLDG 720

Qy        721 GAGDDRLFGGKGNDRLSGDEGDDLLDGGSGDDVLNGGAGNDVYIFRKGDGNDTLYDGTGN 780
```

```
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       721 GAGDDRLFGGKGNDRLSGDEGDDLLDGGSGDDVLNGGAGNDVYIFRKGDGNDTLYDGTGN 780

Qy       781 DKLAFADANISDIMIERTKEGIIVKRNDHSGSINIPRWYITSNLQNYQSNKTDHKIEQLI 840
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       781 DKLAFADANISDIMIERTKEGIIVKRNDHSGSINIPRWYITSNLQNYQSNKTDHKIEQLI 840

Qy       841 GKDGSYITSDQIDKILQDKKDGTVITSQELKKLADENKSQKLSASDIASSLNKLVGSMAL 900
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       841 GKDGSYITSDQIDKILQDKKDGTVITSQELKKLADENKSQKLSASDIASSLNKLVGSMAL 900

Qy       901 FGTANSVSSNALQPITQPTQGILAPSV 927
             |||||||||||||||||||||||||||
Db       901 FGTANSVSSNALQPITQPTQGILAPSV 927


RESULT 2
AAR86998
ID   AAR86998 standard; Protein; 758 AA.
XX
AC   AAR86998;
XX
DT   04-JUL-1996  (first entry)
XX
DE   Enterohaemorrhagic E.coli hlyA gene product.
XX
KW   Enterohaemorrhagic Escherichia coli; virulent; EHEC; O157:H7 serotype;
KW   detection; probe; primer; hlyA gene; enterohaemorrhagic colitis;
KW   haemolytic uremic syndrome; mesenteric adenitis.
XX
OS   Escherichia coli (enterohaemorrhagic).
XX
PN   US5475098-A.
XX
PD   12-DEC-1995.
XX
PF   14-JUN-1994;   94US-0258188.
XX
PR   14-JUN-1994;   94US-0258188.
XX
PA   (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
PI   Hall RH,  Xu JG;
XX
DR   WPI; 1996-048546/05.
DR   N-PSDB; AAT08098.
XX
PT   Entero:haemorrhagic E. coli (EHEC) nucleic acid sequences - useful
PT   for probe and primer design for sensitive and specific detection of
PT   EHEC
XX
PS   Claim 1; Columns 37-42; 32pp; English.
XX
CC   Enterohaemorrhagic E.coli (EHEC) associated with enterohaemorrhagic
CC   colitis, haemolytic uremic syndrome and mesenteric adenitis have
CC   been found to carry a hlyA gene and a hylB gene, separated by an
CC   intergenic region. The hly genes and the intergenic region are
CC   absent from bacteria not associated with these diseases and so
CC   provide a useful target for detecting EHEC pathogens, esp. O157:H7
CC   serotype E.coli. The present sequence is that of the protein
CC   encoded by the EHEC hlyA gene.
XX
SQ   Sequence   758 AA;

  Query Match             1.5%;  Score 14;  DB 17;  Length 758;
  Best Local Similarity   100.0%;  Pred. No. 0.00048;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       379 ISGILEASKQAMFE 392
             ||||||||||||||
Db       152 ISGILEASKQAMFE 165


RESULT 3
AAW22159
ID   AAW22159 standard; Protein; 1049 AA.
XX
AC   AAW22159;
XX
DT   16-FEB-1998  (first entry)
XX
DE   ApxIIIB protein.
XX
KW   RTX toxin; apxICA gene; apxIBD gene; apxIIAB'C gene; apxIIIABCD gene;
KW   repeat in toxins toxin; cell-associated RTX toxin; vaccine production;
KW   therapy; A. pleuropneumoniae infection; swine pleuropneumonia.
XX
OS   Actinobacillus pleuropneumoniae.
XX
PN   CA2170839-A.
XX
PD   02-SEP-1996.
XX
PF   01-MAR-1996;   96CA-2170839.
XX
PR   01-MAR-1995;   95US-0396244.
XX
PA   (UYGU-) UNIV GUELPH.
XX
PI   MacInnes J,  Mallard B,  Ricciatti P,  Rosendal S;
XX
DR   WPI; 1997-245536/23.
DR   N-PSDB; AAT73220.
XX
PT   Preparations of microorganisms producing cell-associated RTX toxins
PT   - especially for production of vaccines against swine
PT   pleuro-pneumonia
XX
PS   Disclosure; Pages 107-110; 151pp; English.
XX
CC   AAW22151-W22161 represent A. pleuropneumoniae RTX (repeat in toxins)
CC   toxins. These sequence are encoded by the apxICA, apxIBD, apxIIAB'C, and
CC   apxIIIABCD genes (see AAT73217-T73220), and can be expressed by
CC   microorganisms used in the preparations of the invention. The
CC   preparations are bacterial preparations comprising one or more isolated
CC   and purified strains of a microorganism that produces one or more RTX
CC   toxins, where the strains have at least one cell-associated RTX toxin.
CC   The preparations are used for production of vaccines for the prophylaxis
CC   and treatment of infectious diseases caused by microorganisms that
CC   produce RTX toxins, where the strains have been attenuated or
CC   inactivated. The vaccines are preferably against Actinobacillus
CC   pleuropneumoniae infection (swine pleuropneumonia). It has been found
CC   that A. pleuropneumonia produces significant quantities of
CC   cell-associated RTX toxins when cultured under certain conditions, and
CC   that the whole-cell protein composition of the cultures corresponds to
CC   the whole-cell protein profiles obtained from cells recovered at
CC   necropsy from the pleural fluid of infected swine. Vaccination with a
CC   bacterin prepared from heat-inactivated cultures having significant
CC   quantities of cell-associated RTX toxins give significant protection of
CC   swine against challenge with homologous strains.
XX
SQ   Sequence   1049 AA;

  Query Match             1.5%;  Score 14;  DB 18;  Length 1049;
  Best Local Similarity   100.0%;  Pred. No. 0.00064;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       241 KVAAGFELSNQVIG 254
             ||||||||||||||
Db       275 KVAAGFELSNQVIG 288


RESULT 4
AAY51412
ID   AAY51412 standard; protein; 1049 AA.
XX
```

```
AC   AAY51412;
XX
DT   05-MAY-2000  (first entry)
XX
DE   A. pleuropneumoniae apxIIIA protein.
XX
KW   RTX toxin; vaccine; Repeats in Toxins; immunostimulatory; antimicrobial;
KW   antiinflammatory; antiarthritic; antiabortive; treatment; pneumonia;
KW   pleuropneumonia; septicemia; nephritis; arthritis; endocarditis;
KW   shipping fever; abortion; whooping cough; sleepy foal disease; joint ill;
KW   urinary infection; peritonitis; meningitis; gastroenteritis;
KW   passive immunization; apxIIIA.
XX
OS   Actinobacillus pleuropneumoniae.
XX
PN   US6019984-A.
XX
PD   01-FEB-2000.
XX
PF   23-DEC-1996;   96US-0772270.
XX
PR   01-MAR-1995;   95US-0396244.
XX
PA   (UYGU-) UNIV GUELPH.
XX
PI   Mallard B,  Rosendal S,  MacInnes J,  Ricciatti P;
XX
DR   WPI; 2000-146864/13.
DR   N-PSDB; AAZ88587.
XX
PT   Bacterial preparation comprising microorganisms which produce a member
PT   of the Repeats in Toxins (RTX) family, useful for treating swine
PT   pleuropneumonia, arthritis in swine, shipping fever and abortion in
PT   cattle, and sleepy foal disease -
XX
PS   Disclosure; Column 77-84; 96pp; English.
XX
CC   This invention describes a novel bacterial preparation (I) which
CC   comprises one or more isolated and purified strain(s) of a microorganism,
CC   cultured in tryptone yeast extract (TYE) broth, which produces one or
CC   more RTX toxins (belonging to the family of toxins referred to as Repeats
CC   in Toxins), where the strain(s) have at least one RTX toxin which is
CC   cell-associated. The products of the invention have immunostimulatory,
CC   antimicrobial, antiinflammatory, antiarthritic and antiabortive activity.
CC   The bacterial preparation may be used as vaccines for the prophylaxis and
CC   treatment of infectious diseases caused by strains of microorganisms
CC   which produce one or more RTX toxins. The infectious diseases are swine
CC   pleuropneumonia, pneumonia, septicemia, nephritis and arthritis in swine;
CC   septicemia, nephritis, endocarditis and arthritis in piglets; shipping
CC   fever and abortion in cattle; whooping cough, sleepy foal disease or
CC   joint ill (purulent nephritis, arthritis) in foals; septicemia,
CC   polyarthritis and abortion in horses; and urinary infections,
CC   peritonitis, meningitis, and gastroenteritis. The bacterial preparations
CC   may also be used to prepare antibodies which may be used as a means of
CC   passive immunization. This sequence represents the Actinobacillus
CC   pleuropneumoniae apxIIIA protein described in the method of the
CC   invention.
XX
SQ   Sequence   1049 AA;

  Query Match             1.5%;  Score 14;  DB 21;  Length 1049;
  Best Local Similarity   100.0%;  Pred. No. 0.00064;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        241 KVAAGFELSNQVIG 254
              ||||||||||||||
Db        275 KVAAGFELSNQVIG 288


RESULT 5
AAR54781
ID   AAR54781 standard; Protein; 1244 AA.
XX
AC   AAR54781;
XX
DT   25-MAR-2003  (updated)
DT   15-OCT-1994  (first entry)
XX
DE   Leukotoxin AppIIIA.
XX
KW   Leukotoxin; AppIIIA; pleuropneumonia; vaccine; diagnostic.
XX
OS   Actinobacillus pleuropneumoniae.
XX
FH   Key             Location/Qualifiers
FT   Domain          327..345
FT                   /label= transmembrane domain
FT   Domain          488..503
FT                   /label= transmembrane domain
FT   Domain          570..587
FT                   /label= transmembrane domain
XX
PN   WO9409821-A1.
XX
PD   11-MAY-1994.
XX
PF   04-NOV-1993;   93WO-US10500.
XX
PR   05-NOV-1992;   92US-0972229.
PR   03-JUN-1993;   93US-0072285.
XX
PA   (CORR ) CORNELL RES FOUND INC.
XX
PI   Chang Y;
XX
DR   WPI; 1994-167130/20.
DR   N-PSDB; AAQ64827.
XX
PT   DNA encoding Actinobacillus pleuropneumoniae leukotoxin - used to
PT   prepare prods. for use in vaccines for porcine pleuropneumonia
PT   and as diagnostic reagents
XX
PS   Disclosure; Fig 3; 65pp; English.
XX
CC   The AppIIIA gene having the sequence given in AAQ64827 was isolated
CC   from phage lambda clones yfc 26-28 and yfc 31-32 of a genomic
CC   library of A. pleuropneumoniae ser. 2.  The gene encodes a novel
CC   leukotoxin (sequence AAR54781).
CC   (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ   Sequence   1244 AA;

  Query Match             1.5%;  Score 14;  DB 15;  Length 1244;
  Best Local Similarity   100.0%;  Pred. No. 0.00075;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        241 KVAAGFELSNQVIG 254
              ||||||||||||||
Db        448 KVAAGFELSNQVIG 461


RESULT 6
AAR12561
ID   AAR12561 standard; Protein; 956 AA.
XX
AC   AAR12561;
XX
DT   25-MAR-2003  (updated)
DT   26-JUL-1991  (first entry)
XX
DE   APPA haemolysin antigen encoded by appA.
XX
KW   Vaccine; porcine pleuroneumonia; haemolysin; cytotoxin.
XX
OS   Actinobacillus pleuroneumoniae serotype 5.
XX
```

```
PN   WO9106653-A.
XX
PD   16-MAY-1991.
XX
PF   31-OCT-1990;   90WO-US06350.
XX
PR   31-OCT-1989;   89US-0429273.
XX
PA   (TEXA ) UNIV TEXAS A & M SYSTEM.
XX
PI   Struck DK,  Young RF,  Chang YF;
XX
DR   WPI; 1991-164202/22.
DR   N-PSDB; AAQ11873.
XX
PT   DNA encoding Actinobacillus pleuropneumoniae haemolysin - used
PT   for producing haemolysin antigen for vaccinating pigs against
PT   porcine pleuro-pneumonia.
XX
PS   Claim 8; Fig 1; 28pp; English.
XX
CC   The sequence is encoded by the appA gene carried on a 3.8 kb
CC   section of clone yfc5 which was isolated by immunological screen-
CC   ing. It forms the appCA antigen with the appC protein encoded on
CC   the same section of DNA.  The appA antigen is not in itself
CC   haemolytic but does elicit an immune response and can be used in
CC   vaccines.
CC   See also AAR12127.
CC   (Updated on 25-MAR-2003 to correct PA field.)
XX
SQ   Sequence   956 AA;

  Query Match             1.2%;  Score 11;  DB 12;  Length 956;
  Best Local Similarity   100.0%;  Pred. No. 0.43;
  Matches   11;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        369 ALLVAGVTGLI 379
              |||||||||||
Db        392 ALLVAGVTGLI 402


RESULT 7
AAW22156
ID   AAW22156 standard; Protein; 956 AA.
XX
AC   AAW22156;
XX
DT   16-FEB-1998  (first entry)
XX
DE   ApxIIC protein.
XX
KW   RTX toxin; apxICA gene; apxIBD gene; apxIIAB'C gene; apxIIIABCD gene;
KW   repeat in toxins toxin; cell-associated RTX toxin; vaccine production;
KW   therapy; A. pleuropneumoniae infection; swine pleuropneumonia.
XX
OS   Actinobacillus pleuropneumoniae.
XX
PN   CA2170839-A.
XX
PD   02-SEP-1996.
XX
PF   01-MAR-1996;   96CA-2170839.
XX
PR   01-MAR-1995;   95US-0396244.
XX
PA   (UYGU-) UNIV GUELPH.
XX
PI   MacInnes J,  Mallard B,  Ricciatti P,  Rosendal S;
XX
DR   WPI; 1997-245536/23.
DR   N-PSDB; AAT73219.
XX
PT   Preparations of microorganisms producing cell-associated RTX toxins
PT   - especially for production of vaccines against swine
PT   pleuro-pneumonia
XX
PS   Disclosure; Pages 101-104; 151pp; English.
XX
CC   AAW22151-W22161 represent A. pleuropneumoniae RTX (repeat in toxins)
CC   toxins. These sequence are encoded by the apxICA, apxIBD, apxIIAB'C, and
CC   apxIIIABCD genes (see AAT73217-T73220), and can be expressed by
CC   microorganisms used in the preparations of the invention. The
CC   preparations are bacterial preparations comprising one or more isolated
CC   and purified strains of a microorganism that produces one or more RTX
CC   toxins, where the strains have at least one cell-associated RTX toxin.
CC   The preparations are used for production of vaccines for the prophylaxis
CC   and treatment of infectious diseases caused by microorganisms that
CC   produce RTX toxins, where the strains have been attenuated or
CC   inactivated. The vaccines are preferably against Actinobacillus
CC   pleuropneumoniae infection (swine pleuropneumonia). It has been found
CC   that A. pleuropneumonia produces significant quantities of
CC   cell-associated RTX toxins when cultured under certain conditions, and
CC   that the whole-cell protein composition of the cultures corresponds to
CC   the whole-cell protein profiles obtained from cells recovered at
CC   necropsy from the pleural fluid of infected swine. Vaccination with a
CC   bacterin prepared from heat-inactivated cultures having significant
CC   quantities of cell-associated RTX toxins give significant protection of
CC   swine against challenge with homologous strains.
XX
SQ   Sequence   956 AA;

  Query Match             1.2%;  Score 11;  DB 18;  Length 956;
  Best Local Similarity   100.0%;  Pred. No. 0.43;
  Matches   11;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        369 ALLVAGVTGLI 379
              |||||||||||
Db        392 ALLVAGVTGLI 402


RESULT 8
AAY51410
ID   AAY51410 standard; protein; 956 AA.
XX
AC   AAY51410;
XX
DT   05-MAY-2000  (first entry)
XX
DE   A. pleuropneumoniae clyIIA protein.
XX
KW   RTX toxin; vaccine; Repeats in Toxins; immunostimulatory; antimicrobial;
KW   antiinflammatory; antiarthritic; antiabortive; treatment; pneumonia;
KW   pleuropneumonia; septicemia; nephritis; arthritis; endocarditis;
KW   shipping fever; abortion; whooping cough; sleepy foal disease; joint ill;
KW   urinary infection; peritonitis; meningitis; gastroenteritis;
KW   passive immunization; clyIIA.
XX
OS   Actinobacillus pleuropneumoniae.
XX
PN   US6019984-A.
XX
PD   01-FEB-2000.
XX
PF   23-DEC-1996;   96US-0772270.
XX
PR   01-MAR-1995;   95US-0396244.
XX
PA   (UYGU-) UNIV GUELPH.
XX
PI   Mallard B,  Rosendal S,  MacInnes J,  Ricciatti P;
XX
DR   WPI; 2000-146864/13.
DR   N-PSDB; AAZ88586.
XX
PT   Bacterial preparation comprising microorganisms which produce a member
PT   of the Repeats in Toxins (RTX) family, useful for treating swine
```

```
PT   pleuropneumonia, arthritis in swine, shipping fever and abortion in
PT   cattle, and sleepy foal disease -
XX
PS   Disclosure; Column 67-72; 96pp; English.
XX
CC   This invention describes a novel bacterial preparation (I) which
CC   comprises one or more isolated and purified strain(s) of a microorganism,
CC   cultured in tryptone yeast extract (TYE) broth, which produces one or
CC   more RTX toxins (belonging to the family of toxins referred to as Repeats
CC   in Toxins), where the strain(s) have at least one RTX toxin which is
CC   cell-associated. The products of the invention have immunostimulatory,
CC   antimicrobial, antiinflammatory, antiarthritic and antiabortive activity.
CC   The bacterial preparation may be used as vaccines for the prophylaxis and
CC   treatment of infectious diseases caused by strains of microorganisms
CC   which produce one or more RTX toxins. The infectious diseases are swine
CC   pleuropneumonia, pneumonia, septicemia, nephritis and arthritis in swine;
CC   septicemia, nephritis, endocarditis and arthritis in piglets; shipping
CC   fever and abortion in cattle; whooping cough, sleepy foal disease or
CC   joint ill (purulent nephritis, arthritis) in foals; septicemia,
CC   polyarthritis and abortion in horses; and urinary infections,
CC   peritonitis, meningitis, and gastroenteritis. The bacterial preparations
CC   may also be used to prepare antibodies which may be used as a means of
CC   passive immunization. This sequence represents the Actinobacillus
CC   pleuropneumoniae clyIIA protein described in the method of the invention.
XX
SQ   Sequence   956 AA;

  Query Match             1.2%;  Score 11;  DB 21;  Length 956;
  Best Local Similarity   100.0%;  Pred. No. 0.43;
  Matches   11;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        369 ALLVAGVTGLI 379
              |||||||||||
Db        392 ALLVAGVTGLI 402


RESULT 9
AAW22152
ID   AAW22152 standard; Protein; 1022 AA.
XX
AC   AAW22152;
XX
DT   16-FEB-1998  (first entry)
XX
DE   ApxIA protein.
XX
KW   RTX toxin; apxICA gene; apxIBD gene; apxIIAB'C gene; apxIIIABCD gene;
KW   repeat in toxins toxin; cell-associated RTX toxin; vaccine production;
KW   therapy; A. pleuropneumoniae infection; swine pleuropneumonia.
XX
OS   Actinobacillus pleuropneumoniae.
XX
PN   CA2170839-A.
XX
PD   02-SEP-1996.
XX
PF   01-MAR-1996;   96CA-2170839.
XX
PR   01-MAR-1995;   95US-0396244.
XX
PA   (UYGU-) UNIV GUELPH.
XX
PI   MacInnes J,  Mallard B,  Ricciatti P,  Rosendal S;
XX
DR   WPI; 1997-245536/23.
DR   N-PSDB; AAT73217.
XX
PT   Preparations of microorganisms producing cell-associated RTX toxins
PT   - especially for production of vaccines against swine
PT   pleuro-pneumonia
XX
PS   Disclosure; Pages 88-91; 151pp; English.
XX
CC   AAW22151-W22161 represent A. pleuropneumoniae RTX (repeat in toxins)
CC   toxins. These sequence are encoded by the apxICA, apxIBD, apxIIAB'C, and
CC   apxIIIABCD genes (see AAT73217-T73220), and can be expressed by
CC   microorganisms used in the preparations of the invention. The
CC   preparations are bacterial preparations comprising one or more isolated
CC   and purified strains of a microorganism that produces one or more RTX
CC   toxins, where the strains have at least one cell-associated RTX toxin.
CC   The preparations are used for production of vaccines for the prophylaxis
CC   and treatment of infectious diseases caused by microorganisms that
CC   produce RTX toxins, where the strains have been attenuated or
CC   inactivated. The vaccines are preferably against Actinobacillus
CC   pleuropneumoniae infection (swine pleuropneumonia). It has been found
CC   that A. pleuropneumonia produces significant quantities of
CC   cell-associated RTX toxins when cultured under certain conditions, and
CC   that the whole-cell protein composition of the cultures corresponds to
CC   the whole-cell protein profiles obtained from cells recovered at
CC   necropsy from the pleural fluid of infected swine. Vaccination with a
CC   bacterin prepared from heat-inactivated cultures having significant
CC   quantities of cell-associated RTX toxins give significant protection of
CC   swine against challenge with homologous strains.
XX
SQ   Sequence   1022 AA;

  Query Match             1.2%;  Score 11;  DB 18;  Length 1022;
  Best Local Similarity   100.0%;  Pred. No. 0.45;
  Matches   11;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        266 AQRVAAGLSTT 276
              |||||||||||
Db        289 AQRVAAGLSTT 299


RESULT 10
AAY51406
ID   AAY51406 standard; protein; 1022 AA.
XX
AC   AAY51406;
XX
DT   05-MAY-2000  (first entry)
XX
DE   A. pleuropneumoniae apxIA protein.
XX
KW   RTX toxin; vaccine; Repeats in Toxins; immunostimulatory; antimicrobial;
KW   antiinflammatory; antiarthritic; antiabortive; treatment; pneumonia;
KW   pleuropneumonia; septicemia; nephritis; arthritis; endocarditis;
KW   shipping fever; abortion; whooping cough; sleepy foal disease; joint ill;
KW   urinary infection; peritonitis; meningitis; gastroenteritis;
KW   passive immunization; apxIA.
XX
OS   Actinobacillus pleuropneumoniae.
XX
PN   US6019984-A.
XX
PD   01-FEB-2000.
XX
PF   23-DEC-1996;   96US-0772270.
XX
PR   01-MAR-1995;   95US-0396244.
XX
PA   (UYGU-) UNIV GUELPH.
XX
PI   Mallard B,  Rosendal S,  MacInnes J,  Ricciatti P;
XX
DR   WPI; 2000-146864/13.
DR   N-PSDB; AAZ88584.
XX
PT   Bacterial preparation comprising microorganisms which produce a member
PT   of the Repeats in Toxins (RTX) family, useful for treating swine
PT   pleuropneumonia, arthritis in swine, shipping fever and abortion in
PT   cattle, and sleepy foal disease -
XX
PS   Disclosure; Column 47-52; 96pp; English.
XX
```

```
CC   This invention describes a novel bacterial preparation (I) which
CC   comprises one or more isolated and purified strain(s) of a microorganism,
CC   cultured in tryptone yeast extract (TYE) broth, which produces one or
CC   more RTX toxins (belonging to the family of toxins referred to as Repeats
CC   in Toxins), where the strain(s) have at least one RTX toxin which is
CC   cell-associated. The products of the invention have immunostimulatory,
CC   antimicrobial, antiinflammatory, antiarthritic and antiabortive activity.
CC   The bacterial preparation may be used as vaccines for the prophylaxis and
CC   treatment of infectious diseases caused by strains of microorganisms
CC   which produce one or more RTX toxins. The infectious diseases are swine
CC   pleuropneumonia, pneumonia, septicemia, nephritis and arthritis in swine;
CC   septicemia, nephritis, endocarditis and arthritis in piglets; shipping
CC   fever and abortion in cattle; whooping cough, sleepy foal disease or
CC   joint ill (purulent nephritis, arthritis) in foals; septicemia,
CC   polyarthritis and abortion in horses; and urinary infections,
CC   peritonitis, meningitis, and gastroenteritis. The bacterial preparations
CC   may also be used to prepare antibodies which may be used as a means of
CC   passive immunization. This sequence represents the Actinobacillus
CC   pleuropneumoniae apxIA protein described in the method of the invention.
XX
SQ   Sequence   1022 AA;

  Query Match             1.2%;  Score 11;  DB 21;  Length 1022;
  Best Local Similarity   100.0%;  Pred. No. 0.45;
  Matches   11;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        266 AQRVAAGLSTT 276
              |||||||||||
Db        289 AQRVAAGLSTT 299


RESULT 11
AAR76991
ID   AAR76991 standard; Protein; 1023 AA.
XX
AC   AAR76991;
XX
DT   13-MAR-1996  (first entry)
XX
DE   LhaA (low homology to appA) protein.
XX
KW   RTX; repeat; vaccine; antigenic; immunisation; pneumonia;
KW   Apl disease.
XX
OS   Actinobacillus pleuropneumoniae.
XX
FH   Key                Location/Qualifiers
FT   misc_difference 750
FT                      /note = "corresponds to ATT codon"
FT   misc_difference 922
FT                      /note = "corresponds to ATT codon"
XX
PN   JP07138185-A.
XX
PD   30-MAY-1995.
XX
PF   23-JUN-1993;   93JP-0152264.
XX
PR   23-JUN-1993;   93JP-0152264.
XX
PA   (NISK ) NIPPON SEIBUTSU KAGAKU KENKYUSHO ZH.
XX
DR   WPI; 1995-228639/30.
DR   N-PSDB; AAQ94783, AAQ94784.
XX
PT   A vaccine contg. a product of the LhaA gene as the active ingredient
PT   - for prophylaxis of Actinobacillus (Haemophilus) pleuropneumoniae
PT   infectious diseases.
XX
PS   Claim 3; Page 8-11; 15pp; Japanese.
XX
CC   AAR76991 represents the lhaA (low homology to appA) gene product.
CC   The lhaA protein (which is a toxin component and contains a repeat
CC   region) and fragments of this protein are useful as the active
CC   ingredient in vaccines for the prophylaxis of Actinobacillus
CC   (Haemophilus) pleuropneumoniae infectious (Apl) diseases.
XX
SQ   Sequence   1023 AA;

  Query Match             1.2%;  Score 11;  DB 16;  Length 1023;
  Best Local Similarity   100.0%;  Pred. No. 0.46;
  Matches   11;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        266 AQRVAAGLSTT 276
              |||||||||||
Db        289 AQRVAAGLSTT 299


RESULT 12
AAE04637
ID   AAE04637 standard; Protein; 450 AA.
XX
AC   AAE04637;
XX
DT   10-SEP-2001  (first entry)
XX
DE   Pasteurella haemolytica modified leukotoxin 50 (lkt50) protein.
XX
KW   Leukotoxin 50; lkt50; respiratory disease; infection; therapy;
KW   immunostimulant; antibacterial; vaccine; transgenic plant;
KW   animal feed; mutant; mutein.
XX
OS   Pasteurella haemolytica.
OS   Synthetic.
XX
PN   WO200144289-A2.
XX
PD   21-JUN-2001.
XX
PF   15-DEC-2000; 2000WO-CA01498.
XX
PR   17-DEC-1999;   99US-0172148.
XX
PA   (UYGU-) UNIV GUELPH OFFICE.
XX
PI   Lo RYC,  Shewen PE,  Lee RWH,  Hodgins D,  Strommer JN;
XX
DR   WPI; 2001-408470/43.
DR   N-PSDB; AAD08976.
XX
PT   Modified leukotoxin polypeptide is useful in a vaccine to prevent or
PT   treat Mannheimia (Pasteurella) infection (particularly M. haemolytica
PT   infection), and disease associated with a leukotoxin, e.g., respiratory
PT   disease -
XX
PS   Claim 4; Fig 12; 70pp; English.
XX
CC   The present sequence is Pasteurella (Mannheimia) haemolytica
CC   modified leukotoxin-50 (lkt50) protein. The modification comprises
CC   the removal of amino acids within the hydrophobic transmembrane
CC   domain of a full length leukotoxin protein. Modified leukotoxin
CC   sequences are used in vaccines to treat or prevent diseases associated
CC   with leukotoxin, e.g., respiratory disease, and Mannheimia infection
CC   (particularly M. haemolytica infection). In addition, the vaccine
CC   is used to prepare a medicament. Furthermore, the plant transformed
CC   with modified leukotoxin sequences is fed to an animal such as a
CC   ruminant, to prevent or treat respiratory diseases.
XX
SQ   Sequence   450 AA;

  Query Match             1.1%;  Score 10;  DB 22;  Length 450;
  Best Local Similarity   100.0%;  Pred. No. 1.9;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        443 AERVIAITQQ 452
              ||||||||||
```

```
Db         11 AERVIAITQQ 20


RESULT 13
AAY33929
ID   AAY33929 standard; Protein; 490 AA.
XX
AC   AAY33929;
XX
DT   09-NOV-1999  (first entry)
XX
DE   Leukotoxin carrier protein (LKT 114).
XX
KW   Myostatin; mouse; rabbit; human; baboon; bovine; porcine; ovine; chick;
KW   turkey; zebrafish; immune response; vaccine; body weight; muscle mass;
KW   mammary gland tissue; lactation; feed uptake; muscle degeneration;
KW   GDF11 activity; LKT 114; leukotoxin carrier protein.
XX
OS   Pasteurella haemolytica.
XX
PN   WO9942573-A1.
XX
PD   26-AUG-1999.
XX
PF   19-FEB-1999;   99WO-CA00128.
XX
PR   19-FEB-1998;   98US-0075213.
XX
PA   (BIOS-) BIOSTAR INC.
XX
PI   Barker CA,  Morsey M;
XX
DR   WPI; 1999-527471/44.
DR   N-PSDB; AAX99361.
XX
PT   New myostatin peptide, multimers and immunoconjugates for eliciting
PT   an immune response in a vertebrate against a myostatin immunogen
XX
PS   Example 2; Fig 15; 109pp; English.
XX
CC   The invention provides myostatin peptides consisting of 3-100 amino
CC   acids, derived from a region of mouse, rabbit, human, baboon, bovine,
CC   porcine, ovine, chick, turkey or zebrafish myostatin (see sequences
CC   AAY33930-939). The myostatin peptides are derived preferably from a
CC   region of amino acid residues 1-275, 25-300, 50-325 or 75-350 of the
CC   above sequences. The peptides and the nucleic acids encoding the peptides
CC   are useful as vaccines for eliciting an immune response in a vertebrate
CC   against a myostatin immunogen. They result in increasing body weight,
CC   muscle mass, number and size of muscle cells, muscle strength, mammary
CC   gland tissue, lactation, appetite or feed uptake, life span of the
CC   vertebrate, and cause a reduction in body fat content, useful for muscle
CC   wasting conditions. The vaccines are also useful for treating a disorder
CC   which comprises degeneration or wasting of muscle in a vertebrate, and
CC   useful for modulating GDF11 activity. The present sequence represents a
CC   leukotoxin carrier protein (LKT 114) which is used as an immunological
CC   carrier protein to form a myostatin immunoconjugate.
XX
SQ   Sequence   490 AA;

  Query Match             1.1%;  Score 10;  DB 20;  Length 490;
  Best Local Similarity   100.0%;  Pred. No. 2.1;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        265 LAQRVAAGLS 274
              ||||||||||
Db        256 LAQRVAAGLS 265


RESULT 14
AAW03943
ID   AAW03943 standard; Protein; 544 AA.
XX
AC   AAW03943;
XX
DT   20-NOV-1996  (first entry)
XX
DE   LKT-GnRH protein fusion from pCB111.
XX
KW   Leukotoxin; LKT; gonadotropin-releasing hormone; GnRH;
KW   fusion protein; immunogen; vaccine; fertility control;
KW   contraceptive; sterilisation.
XX
OS   Chimeric Pasteurella haemolytica A1 strain B122;
OS   Chimeric synthetic.
XX
FH   Key             Location/Qualifiers
FT   Domain          1..493
FT                   /label= LKT
FT   Domain          494..544
FT                   /label= GnRH_repeat_domain
XX
PN   WO9624675-A1.
XX
PD   15-AUG-1996.
XX
PF   24-JAN-1996;   96WO-CA00049.
XX
PR   10-FEB-1995;   95US-0387156.
XX
PA   (UYSA-) UNIV SASKATCHEWAN.
XX
PI   Manns JG,  Potter AA;
XX
DR   WPI; 1996-384447/38.
DR   N-PSDB; AAT37177.
XX
PT   Gonadotropin-releasing hormone multimer fusion proteins - with
PT   leukotoxin polypeptide for increased immunogenicity, useful in
PT   antifertility vaccine prodn.
XX
PS   Claim 8; Fig 7A-7E; 87pp; English.
XX
CC   A chimeric protein (AAW03943) is composed of a fusion between
CC   a truncated leukotoxin (LKT-111) from Pasteurella haemolytica
CC   and a 4-copy gonadoliberin-releasing hormone (GnRH) repeat
CC   sequence (see also AAW03944).  It is the product of a chimeric
CC   gene (AAT37177) produced by deleting an approx. 1300 bp sequence
CC   from pCB113 (see also AAT37176) coding for amino acids 352-784
CC   of LKT-352.  Recombinant plasmid pCB111 (LKT 111:4 copy GnRH,
CC   ATCC 69748) was obtd.  Escherichia coli transformants produced
CC   the chimeric protein, which is useful as a vaccine for fertility
CC   control, esp. immunological sterilisation of domestic or farm
CC   animals.
XX
SQ   Sequence   544 AA;

  Query Match             1.1%;  Score 10;  DB 17;  Length 544;
  Best Local Similarity   100.0%;  Pred. No. 2.3;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        265 LAQRVAAGLS 274
              ||||||||||
Db        256 LAQRVAAGLS 265


RESULT 15
AAW79570
ID   AAW79570 standard; Protein; 544 AA.
XX
AC   AAW79570;
XX
DT   24-DEC-1998  (first entry)
XX
DE   LKT-GnRH chimeric protein.
XX
KW   Chimera; pCB111; LKT 111; GnRH; Gonadotropin releasing hormone; multimer;
```

```
KW   cytotoxic activity; antigen presentation; immune response; vaccine;
KW   tumour.
XX
OS   Synthetic.
XX
PN   WO9806848-A1.
XX
PD   19-FEB-1998.
XX
PF   08-AUG-1997;   97WO-CA00559.
XX
PR   09-AUG-1996;   96US-0694865.
XX
PA   (UYSA-) UNIV SASKATCHEWAN.
XX
PI   Manns JG,  Potter AA;
XX
DR   WPI; 1998-159540/14.
DR   N-PSDB; AAV61532.
XX
PT   Chimeric protein of leukotoxin and gonadotropin releasing hormone  -
PT   useful for, e.g. preparation of vaccines for reduction of incidence
PT   of mammary tumours in mammals
XX
PS   Disclosure; Figure 7.1-5; 118pp; English.
XX
CC   The present sequence represents the LKT-GnRH chimeric protein from
CC   pCB111.  This plasmid contains the LKT 111 polypeptide fused to
CC   four copies of the GnRH peptide.  This chimera lacks cytotoxic activity
CC   which enables there to be an increase in antigen presentation and thus an
CC   optimal immune response.  The removal of this region also enables the
CC   truncated LKT to be expressed at much higher levels and allows the amount
CC   of antigen administered to be reduced.  This chimeric protein comprises a
CC   leukotoxin polypeptide, several multimers, and a GnRH sequence.  The
CC   chimeric protein can be used as a vaccine to help reduce the incidence of
CC   mammary tumours in a mammalian individual.
XX
SQ   Sequence   544 AA;

  Query Match              1.1%;  Score 10;  DB 19;  Length 544;
  Best Local Similarity   100.0%;  Pred. No. 2.3;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        265 LAQRVAAGLS 274
              ||||||||||
Db        256 LAQRVAAGLS 265


RESULT 16
AAE04636
ID   AAE04636 standard; Protein; 608 AA.
XX
AC   AAE04636;
XX
DT   10-SEP-2001  (first entry)
XX
DE   Pasteurella haemolytica modified leukotoxin 66 (lkt66) protein.
XX
KW   Leukotoxin 66; lkt66; respiratory disease; infection; therapy;
KW   immunostimulant; antibacterial; vaccine; transgenic plant;
KW   animal feed; mutant; mutein.
XX
OS   Pasteurella haemolytica.
OS   Synthetic.
XX
PN   WO200144289-A2.
XX
PD   21-JUN-2001.
XX
PF   15-DEC-2000; 2000WO-CA01498.
XX
PR   17-DEC-1999;   99US-0172148.
XX
PA   (UYGU-) UNIV GUELPH OFFICE.
XX
PI   Lo RYC,  Shewen PE,  Lee RWH,  Hodgins D,  Strommer JN;
XX
DR   WPI; 2001-408470/43.
DR   N-PSDB; AAD08975. .
XX
PT   Modified leukotoxin polypeptide is useful in a vaccine to prevent or
PT   treat Mannheimia (Pasteurella) infection (particularly M. haemolytica
PT   infection), and disease associated with a leukotoxin, e.g., respiratory
PT   disease -
XX
PS   Claim 2; Fig 2; 70pp; English.
XX
CC   The present sequence is Pasteurella (Mannheimia) haemolytica
CC   modified leukotoxin-66 (lkt66) protein. The modification comprises
CC   the removal of amino acids within the hydrophobic transmembrane
CC   domain of a full length leukotoxin protein. Modified leukotoxin
CC   sequences are used in vaccines to treat or prevent diseases associated
CC   with leukotoxin, e.g., respiratory disease, and Mannheimia infection
CC   (particularly M. haemolytica infection). In addition, the vaccine
CC   is used to prepare a medicament. Furthermore, the plant transformed
CC   with modified leukotoxin sequences is fed to an animal such as a
CC   ruminant, to prevent or treat respiratory diseases.
XX
SQ   Sequence   608 AA;

  Query Match              1.1%;  Score 10;  DB 22;  Length 608;
  Best Local Similarity   100.0%;  Pred. No. 2.5;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        443 AERVIAITQQ 452
              ||||||||||
Db        116 AERVIAITQQ 125


RESULT 17
AAW79573
ID   AAW79573 standard; Protein; 695 AA.
XX
AC   AAW79573;
XX
DT   24-DEC-1998  (first entry)
XX
DE   LKT-GnRH chimeric protein.
XX
KW   Chimera; pCB122; LKT 111; GnRH; Gonadotropin releasing hormone; multimer;
KW   cytotoxic activity; antigen presentation; immune response; vaccine;
KW   tumour.
XX
OS   Synthetic.
XX
PN   WO9806848-A1.
XX
PD   19-FEB-1998.
XX
PF   08-AUG-1997;   97WO-CA00559.
XX
PR   09-AUG-1996;   96US-0694865.
XX
PA   (UYSA-) UNIV SASKATCHEWAN.
XX
PI   Manns JG,  Potter AA;
XX
DR   WPI; 1998-159540/14.
DR   N-PSDB; AAV61535.
XX
PT   Chimeric protein of leukotoxin and gonadotropin releasing hormone  -
PT   useful for, e.g. preparation of vaccines for reduction of incidence
PT   of mammary tumours in mammals
XX
PS   Claim 9; Figure 9.1-6; 118pp; English.
XX
```

```
CC   The present sequence represents the LKT-GnRH chimeric protein from
CC   pCB122.  This plasmid contains the LKT 111 polypeptide fused to sixteen
CC   copies of the GnRH peptide, in the pattern of: 8 copies of GnRH-LKT 111-8
CC   copies of GnRH.  This chimera lacks cytotoxic activity which enables
CC   there to be an increase in antigen presentation and thus an optimal
CC   immune response.  The removal of this region also enables the truncated
CC   LKT to be expressed at much higher levels and allows the amount of
CC   antigen administered to be reduced.  This chimeric protein comprises a
CC   leukotoxin polypeptide, several multimers, and a GnRH sequence.  The
CC   chimeric protein can be used as a vaccine to help reduce the incidence of
CC   mammary tumours in a mammalian individual.
XX
SQ   Sequence   695 AA;

  Query Match             1.1%;  Score 10;  DB 19;  Length 695;
  Best Local Similarity   100.0%;  Pred. No. 2.9;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        265 LAQRVAAGLS 274
              ||||||||||
Db        359 LAQRVAAGLS 368


RESULT 18
AAY58361
ID   AAY58361 standard; Protein; 695 AA.
XX
AC   AAY58361;
XX
DT   27-MAR-2000  (first entry)
XX
DE   Leukotoxin/gonadotropin-releasing hormone fusion protein.
XX
KW   Leukotoxin; gonadotropin-releasing hormone; GnRH; immunosterilisation;
KW   immunocontraception; vaccine; feline; canine; equine; cervine.
XX
OS   Chimeric - Pasteurella haemolytica.
OS   Chimeric - Mammalia.
XX
PN   WO9962545-A2.
XX
PD   09-DEC-1999.
XX
PF   28-MAY-1999;   99WO-CA00493.
XX
PR   04-JUN-1998;   98US-0088024.
PR   06-MAY-1999;   99US-0306689.
XX
PA   (BIOS-) BIOSTAR INC.
XX
PI   Robbins SC;
XX
DR   WPI; 2000-086857/07.
DR   N-PSDB; AAZ55700.
XX
PT   Hormone immunogens, analogues or antibodies used to manufacture
PT   vaccines for suppression of reproductive behavior and fertility in
PT   vertebrates -
XX
PS   Claim 20; Fig 6A-6F; 88pp; English.
XX
CC   This sequence represents a fusion protein comprising gonadotropin-
CC   releasing hormone (GnRH) immunogens and a Pasteurella
CC   haemolytica leukotoxin (LKT) protein. The fusion protein comprises, in
CC   the N to C-terminal direction, a synthetic peptide sequence (AAY58364),
CC   an eight copy GnRH multimer (composed of two copies of the 4xGnRH
CC   multimer sequence of AAY58363), the LKT protein (which functions as a
CC   carrier protein), and a second eight copy GnRH multimer. The fusion
CC   protein may be used in a vaccine composition for prepubertal
CC   administration to a vertebrate subject to result in prolonged suppression
CC   of reproductive behaviour and/or fertility. GnRH immunogens, analogues or
CC   antibodies that cross-react with endogenous GnRH of a vertebrate subject
CC   are used to manufacture a composition or vaccine for immunosterilisation
CC   or immunocontraception of feline, canine, equine or cervine subjects.
CC   The vaccines are used to suppress reproductive behaviour and/or
CC   fertility for at least 10 months. The prepubertal administration
CC   results in a prolonged, long-term suppression of testicular development
CC   and/or function in males, or a prolonged, long-term suppression of
CC   ovarian development and/or function in females. The methods provide a
CC   viable and desirable alternative to surgical forms of sterilisation that
CC   are currently used.
XX
SQ   Sequence   695 AA;

  Query Match             1.1%;  Score 10;  DB 21;  Length 695;
  Best Local Similarity   100.0%;  Pred. No. 2.9;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        265 LAQRVAAGLS 274
              ||||||||||
Db        359 LAQRVAAGLS 368


RESULT 19
AAY58133
ID   AAY58133 standard; Protein; 695 AA.
XX
AC   AAY58133;
XX
DT   07-MAR-2000  (first entry)
XX
DE   Gonadotropin releasing hormone-leukotoxin fusion protein.
XX
KW   Gonadotropin releasing hormone; GnRH; leukotoxin; LKT; fusion protein;
KW   antibody; immunogenic; chimeric; vaccine; testosterone; androgenic;
KW   non-androgenic; steroid; reduction; weight gain; muscle distribution;
KW   fat distribution; male pattern; boar taint; flavour; impairment;
KW   reliable; immunocastration; meat production.
XX
OS   Chimeric - Mammalia.
OS   Chimeric - Pasteurella haemolytica.
XX
PN   WO9956771-A2.
XX
PD   11-NOV-1999.
XX
PF   05-MAY-1999;   99WO-CA00360.
XX
PR   05-MAY-1998;   98US-0084217.
XX
PA   (BIOS-) BIOSTAR INC.
XX
PI   Manns JG,  Acres SD,  Harland R;
XX
DR   WPI; 2000-062125/05.
DR   N-PSDB; AAZ46400.
XX
PT   Production of uncastrated male food animals using vaccines -
XX
PS   Claim 22; Fig 3A-3F; 87pp; English.
XX
CC   This sequence represents a chimeric gonadotropin releasing
CC   hormone (GnRH)-leukotoxin (LKT) fusion protein, which may be
CC   used as a vaccine. The LKT portion of the protein acts to enhance
CC   the immunogenicity of the multimeric GnRH portion (AAY58135). The
CC   invention relates to a method of using two GnRH immunogen vaccines to
CC   produce uncastrated male animals for meat production, one vaccination
CC   prior to or during the fattening period to reduce circulating
CC   testosterone levels, and the second vaccination about 2-8 weeks before
CC   slaughter to substantially reduce androgenic and/or non-androgenic
CC   steroids. The invention is used to produce food animals that exhibit the
CC   weight gain and muscle/fat distribution of male animals without the
CC   problems associated with male animals. Such problems include "boar
CC   taint", a urine-like odour found in cooked meat of uncastrated pigs which
CC   is caused by steroids stored in the tissues, and similar flavour
CC   impairments in the meat of other intact male animals. The invention is
```

```
CC   more reliable than prior art immunocastration techniques.
XX
SQ   Sequence   695 AA;

  Query Match             1.1%;  Score 10;  DB 21;  Length 695;
  Best Local Similarity   100.0%;  Pred. No. 2.9;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        265 LAQRVAAGLS 274
              ||||||||||
Db        359 LAQRVAAGLS 368


RESULT 20
ABB67818
ID   ABB67818 standard; Protein; 802 AA.
XX
AC   ABB67818;
XX
DT   26-MAR-2002  (first entry)
XX
DE   Drosophila melanogaster polypeptide SEQ ID NO 30246.
XX
KW   Drosophila; developmental biology; cell signalling; insecticide;
KW   pharmaceutical.
XX
OS   Drosophila melanogaster.
XX
PN   WO200171042-A2.
XX
PD   27-SEP-2001.
XX
PF   23-MAR-2001; 2001WO-US09231.
XX
PR   23-MAR-2000; 2000US-191637P.
PR   11-JUL-2000; 2000US-0614150.
XX
PA   (PEKE ) PE CORP NY.
XX
PI   Venter JC,  Adams M,  Li PWD,  Myers EW;
XX
DR   WPI; 2001-656860/75.
DR   N-PSDB; ABL11921.
XX
PT   New isolated nucleic acid detection reagent for detecting 1000 or more
PT   genes from Drosophila and for elucidating cell signalling and cell-cell
PT   interactions -
XX
PS   Disclosure; SEQ ID NO 30246; 21pp + Sequence Listing; English.
XX
CC   The invention relates to an isolated nucleic acid detection reagent
CC   capable of detecting 1000 or more genes from Drosophila. The invention is
CC   useful in developmental biology and in elucidating cell signalling and
CC   cell-cell interactions in higher eukaryotes for the development of
CC   insecticides, therapeutics and pharmaceutical drugs. The invention
CC   discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC   sequences (ABL01840-ABL16175) and the encoded proteins
CC   (ABB57737-ABB72072).
CC   The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence   802 AA;

  Query Match             1.1%;  Score 10;  DB 22;  Length 802;
  Best Local Similarity   100.0%;  Pred. No. 3.3;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        353 AGVSAAAVGS 362
              ||||||||||
Db        484 AGVSAAAVGS 493


RESULT 21
AAR10889
ID   AAR10889 standard; Protein; 924 AA.
XX
AC   AAR10889;
XX
DT   11-APR-1991  (first entry)
XX
DE   Leukotoxin 352 encoded by plasmid pAA352.
XX
KW   LKT; vaccine; antigen; respiratory disease; shipping fever pneumonia.
XX
OS   Pasteurella haemolytica A1 strain B122.
XX
PN   CA2014033-A.
XX
PD   07-OCT-1990.
XX
PF   06-APR-1990;   90CA-2014033.
XX
PR   07-APR-1989;   89US-0335018.
XX
PA   (UYSA-) UNIV SASKATCHEWAN.
XX
PI   Acres SD,  Babiuk LA,  Potter AA,  Lawman MJP;
XX
DR   WPI; 1991-000097/01.
DR   N-PSDB; AAQ10272.
XX
PT   Pasteurella haemolytica proteins and genes - used for producing
PT   vaccines to protect animals esp. cattle from respiratory diseases
PT   e.g. pneumonia.
XX
PS   Claim 13; Fig 5; 87pp; English.
XX
CC   Plasmid pAA352 is derived from pAA114, a clone isolated from a
CC   genomic library of P. haemolytica. The protein, designated "new
CC   leukotoxin" or "LKT 352"  is 98% homologous to authentic leukotoxin.
CC   LKT 352 and pref. antigenic fragments of it,  can be used in
CC   vaccines to protect cattle from respiratory diseases. They can also
CC   be used to produce antibodies for immunoaffinity purificn. of
CC   further proteins. [Fig. contg. sequence v. poor].
CC   See also AAR10890, AAR10909, AAR10910 and AAQ10783.
XX
SQ   Sequence   924 AA;

  Query Match             1.1%;  Score 10;  DB 12;  Length 924;
  Best Local Similarity   100.0%;  Pred. No. 3.7;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        265 LAQRVAAGLS 274
              ||||||||||
Db        254 LAQRVAAGLS 263


RESULT 22
AAR42385
ID   AAR42385 standard; Protein; 924 AA.
XX
AC   AAR42385;
XX
DT   25-MAR-2003  (updated)
DT   19-APR-1994  (first entry)
XX
DE   Recombinant leukotoxin peptide from plasmid pCRR28.
XX
KW   Haemophilus somnus; immunogenic; haemolysin; LppB; LppC;
KW   thromboembolic meningoencephalitis; septicaemia; arthritis;
KW   pneumonia; lktA gene; haemin-binding protein; fusion protein.
XX
OS   Pasteurella haemolytica.
XX
PN   WO9321323-A1.
```

```
XX
PD   28-OCT-1993.
XX
PF   05-APR-1993;   93WO-CA00135.
XX
PR   09-APR-1992;   92US-0865050.
PR   04-JUN-1992;   92US-0893424.
PR   04-JUN-1992;   92US-0893426.
PR   29-MAR-1993;   93US-0038287.
PR   29-MAR-1993;   93US-0038288.
PR   29-MAR-1993;   93US-0038719.
XX
PA   (UYSA-) UNIV SASKATCHEWAN.
XX
PI   Harland RJ,  Pfeiffer CG,  Pontarollo RA,  Potter AA;
PI    Rioux C,  Theisen M;
XX
DR   WPI; 1993-351733/44.
DR   N-PSDB; AAQ51086.
XX
PT   Haemophilus somnus immunogenic proteins used in vaccines -
PT   selected from haemin-binding protein, haemolysin, LppB and LppC,
PT   and corresp. DNA
XX
PS   Disclosure; Fig 11; 119pp; English.
XX
CC   The lppB gene protein was expressed in E. coli as a fusion to the
CC   Pasteurella haemolytica leukotoxin gene lktA coded for by plasmid
CC   pAA352.  The llpB gene fragment was taken from pMS11.  LppB can be
CC   used in vaccines for preventing or treating H. somnus infections,
CC   which cause thromboembolic meningo-encephalitis, septicaemia, arthritis
CC   and pneumonia in vertebrates.
CC   See also AAR42370-86.
CC   (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ   Sequence   924 AA;

  Query Match              1.1%;  Score 10;  DB 14;  Length 924;
  Best Local Similarity   100.0%;  Pred. No. 3.7;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        265 LAQRVAAGLS 274
              ||||||||||
Db        256 LAQRVAAGLS 265

RESULT 23
AAR42380
ID   AAR42380 standard; Protein; 924 AA.
XX
AC   AAR42380;
XX
DT   25-MAR-2003  (updated)
DT   19-APR-1994  (first entry)
XX
DE   Recombinant leukotoxin peptide (split) from plasmid pGCH4.
XX
KW   Haemophilus somnus; immunogenic; haemolysin; LppB; LppC;
KW   thromboembolic meningoencephalitis; septicaemia; arthritis;
KW   pneumonia; lktA gene; haemin-binding protein; fusion protein.
XX
OS   Pasteurella haemolytica.
XX
PN   WO9321323-A1.
XX
PD   28-OCT-1993.
XX
PF   05-APR-1993;   93WO-CA00135.
XX
PR   09-APR-1992;   92US-0865050.
PR   04-JUN-1992;   92US-0893424.
PR   04-JUN-1992;   92US-0893426.
PR   29-MAR-1993;   93US-0038287.
PR   29-MAR-1993;   93US-0038288.
PR   29-MAR-1993;   93US-0038719.
XX
PA   (UYSA-) UNIV SASKATCHEWAN.
XX
PI   Harland RJ,  Pfeiffer CG,  Pontarollo RA,  Potter AA;
PI    Rioux C,  Theisen M;
XX
DR   WPI; 1993-351733/44.
DR   N-PSDB; AAQ51082.
XX
PT   Haemophilus somnus immunogenic proteins used in vaccines -
PT   selected from haemin-binding protein, haemolysin, LppB and LppC,
PT   and corresp. DNA
XX
PS   Disclosure; Fig 6; 119pp; English.
XX
CC   The hmb gene encoding the haemin-binding protein was expressed in
CC   E. coli as a fusion to the Pasteurella haemolytica leukotoxin gene
CC   lktA coded for by plasmid pAA352.  The hmb gene fragment was taken
CC   from pRAP504 and starts at the codon for the 33 rd amino acid residue
CC   of ORF1.  The haemin binding protein can be used in vaccines for
CC   preventing or treating H. somnus infections, which cause thromboembolic
CC   meningo-encephalitis, septicaemia, arthritis and pneumonia in
CC   vertebrates.
CC   See also AAR42370-86.
CC   (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ   Sequence   924 AA;

  Query Match              1.1%;  Score 10;  DB 14;  Length 924;
  Best Local Similarity   100.0%;  Pred. No. 3.7;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        265 LAQRVAAGLS 274
              ||||||||||
Db        256 LAQRVAAGLS 265

RESULT 24
AAR42378
ID   AAR42378 standard; Protein; 924 AA.
XX
AC   AAR42378;
XX
DT   25-MAR-2003  (updated)
DT   19-APR-1994  (first entry)
XX
DE   Recombinant leukotoxin peptide (split) from plasmid pGCH5.
XX
KW   Haemophilus somnus; immunogenic; haemolysin; LppB; LppC;
KW   thromboembolic meningoencephalitis; septicaemia; arthritis;
KW   pneumonia; lktA gene; haemin-binding protein; fusion protein.
XX
OS   Pasteurella haemolytica.
XX
PN   WO9321323-A1.
XX
PD   28-OCT-1993.
XX
PF   05-APR-1993;   93WO-CA00135.
XX
PR   09-APR-1992;   92US-0865050.
PR   04-JUN-1992;   92US-0893424.
PR   04-JUN-1992;   92US-0893426.
PR   29-MAR-1993;   93US-0038287.
PR   29-MAR-1993;   93US-0038288.
PR   29-MAR-1993;   93US-0038719.
XX
PA   (UYSA-) UNIV SASKATCHEWAN.
XX
PI   Harland RJ,  Pfeiffer CG,  Pontarollo RA,  Potter AA;
PI    Rioux C,  Theisen M;
```

```
XX
DR   WPI; 1993-351733/44.
DR   N-PSDB; AAQ51081.
XX
PT   Haemophilus somnus immunogenic proteins used in vaccines -
PT   selected from haemin-binding protein, haemolysin, LppB and LppC,
PT   and corresp. DNA
XX
PS   Disclosure; Fig 5; 119pp; English.
XX
CC   The hmb gene encoding the haemin-binding protein was expressed in
CC   E. coli as a fusion to the Pasteurella haemolytica leukotoxin gene
CC   lktA coded for by plasmid pAA352.  The hmb gene fragment was taken
CC   from pRAP501 and starts at the codon for the third amino acid residue
CC   of ORF1.  The haemin binding protein can be used in vaccines for
CC   preventing or treating H. somnus infections, which cause thromboembolic
CC   meningo-encephalitis, septicaemia, arthritis and pneumonia in
CC   vertebrates.
CC   See also AAR42370-86.
CC   (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ   Sequence   924 AA;

  Query Match             1.1%;  Score 10;  DB 14;  Length 924;
  Best Local Similarity   100.0%;  Pred. No. 3.7;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        265 LAQRVAAGLS 274
              ||||||||||
Db        256 LAQRVAAGLS 265


RESULT 25
AAR14482
ID   AAR14482 standard; Protein; 926 AA.
XX
AC   AAR14482;
XX
DT   15-JAN-1992  (first entry)
XX
DE   LKT352.
XX
KW   Antigen; leukotoxin; vaccine; lktA.
XX
OS   Pasteurella haemolytica.
XX
PN   WO9115237-A.
XX
PD   17-OCT-1991.
XX
PF   17-OCT-1991;   91WO-CA00170.
XX
PR   05-APR-1990;   90US-0504850.
XX
PA   (UYSA-) UNIV SASKATCHEWAN.
XX
PI   Acres SD,  Bariuk LA,  Potter AA,  Lawman MJP;
XX
DR   WPI; 1991-324967/44.
XX
PT   Vaccines for Pasteurella haemolytica infection in cattle -
PT   comprise sub-unit antigens from P haemolytica fimbrial protein,
PT   plasmin receptor, 50 K outer membrane protein and leukotoxin.
XX
PS   Disclosure; Fig 5; 92pp; English.
XX
CC   LKT352 is 98% homologous with authentic leukotoxin and migrates
CC   to the same position on gels.
CC   The LKT352 gene was prepd. as follows: lktA, an MaeI fragment
CC   contg. the gene was ligated into the SmaI site of pUC13 to form
CC   pAA179.  From this, two constructs were made in the ptac-based
CC   vector, pGH432:lacI digested with SmaI.  One, pAA342, consisted of
CC   the 5' AhaIII fragment from lktA while the other, pAA345, contained
CC   the entire MaeI fragment.  Clone pAA342 expressed a truncated
CC   leukotoxin peptide at high levels while pAA345 expressed full
CC   length leukotoxin at very low levels.  The 3' end of the lktA gene
CC   of pAA345 was therefore ligated to StyI/BamHI digested pAA342 to
CC   yield pAA352 contg. the LKT352 sequence.  The protein expressed
CC   from the vector can be used to prepare a subunit vaccine with
CC   other P. haemolytica antigens, e.g. fimbrial protein, plasmin
CC   receptor or 50K outer membrane protein.  The vaccines can be used
CC   to protect cattle from respiratory diseases such as pneumonia, esp.
CC   shipping fever pneumonia.
CC   See also AAR14481, 83,84 and 85.
XX
SQ   Sequence   926 AA;

  Query Match             1.1%;  Score 10;  DB 12;  Length 926;
  Best Local Similarity   100.0%;  Pred. No. 3.7;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        265 LAQRVAAGLS 274
              ||||||||||
Db        256 LAQRVAAGLS 265


RESULT 26
AAR34545
ID   AAR34545 standard; Protein; 926 AA.
XX
AC   AAR34545;
XX
DT   25-MAR-2003  (updated)
DT   23-AUG-1993  (first entry)
XX
DE   Leukotoxin 352 produced from pAA352.
XX
KW   Vector; LKT 352; flanking; recombinant; antigen; somatostatin;
KW   gonadotropin releasing hormone; rotavirus viral protein 4;
KW   carrier protein; lactation; reproduction.
XX
OS   Pasteurella haemolytica.
XX
PN   WO9308290-A1.
XX
PD   29-APR-1993.
XX
PF   15-OCT-1992;   92WO-CA00449.
XX
PR   16-OCT-1991;   91US-0779171.
PR   14-OCT-1992;   92US-0960932.
XX
PA   (UYSA-) UNIV SASKATCHEWAN.
XX
PI   Hughes HPA,  Potter AA,  Redmond MJ;
XX
DR   WPI; 1993-152482/18.
DR   N-PSDB; AAQ41317.
XX
PT   Immunological carrier system with enhanced immunogenicity -
PT   comprises chimeric protein comprising leuco:toxin peptide or
PT   homologous protein fused to antigen esp. somatostatin or
PT   gonadotropin releasing hormone
XX
PS   Disclosure; Fig 3; 95pp; English.
XX
CC   Gene libraries of P. haemolytica A1 (strain B122) were constructed
CC   in lambda gt11 and pUC13.  Resulting clones were used to transform E.
CC   coli and individual colonies were pooled and screened for reaction
CC   with serum from a calf which had survived a P. haemolytica infection
CC   and that had been boosted with a conc. culture supernatant of P.
CC   haemolytica to increase anti-leukotoxin antibody levels.  Positive
CC   colonies were screened for their ability to produce leukotoxin by
CC   incubating cell lysates with bovine neutrophils and measuring the
CC   release of lactate dehydrogenase from the neutrophils.  A 4kb
CC   fragment was obtd.  Progressively larger clones were isolated by
```

```
CC   chromosome walking to isolate full length recombinants of ca. 8kb,
CC   in pAA114.  The clone was subjected to restriction enzyme digestion
CC   to yield two clones, one expressing truncated leukotoxin peptide at
CC   high levels and the other expressing the full length leukotoxin at
CC   low levels.  The 3' end of the lktA gene from the full length clone
CC   was ligated to the truncated gene clone to yield plasmid pAA352.  The
CC   clone was used to produce chimeric proteins by gene fusion with an
CC   antigen coding sequence, e.g. the coding sequence of somatostatin,
CC   gonadotrophin releasing hormone or rotavirus viral protein 4, i.e.
CC   leukotoxin works as a carrier protein to bring about a larger
CC   immune response than the antigen alone.  Immunisation with these
CC   antigens can regulate growth rate, lactation and reproductive
CC   efficiency. See also AAR34546-8.
CC   (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ   Sequence   926 AA;

  Query Match             1.1%;  Score 10;  DB 14;  Length 926;
  Best Local Similarity   100.0%;  Pred. No. 3.7;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        265 LAQRVAAGLS 274
              ||||||||||
Db        256 LAQRVAAGLS 265


RESULT 27
AAR50291
ID   AAR50291 standard; Protein; 926 AA.
XX
AC   AAR50291;
XX
DT   06-OCT-1994  (first entry)
XX
DE   Recombinant leukotoxin from plasmid pAA352.
XX
KW   Vaccine; outer membrane protein; OMP; Haemophilus somnus;
KW   iron regulated protein; leukotoxin; Pasteurella haemolytica;
KW   LKT352.
XX
OS   Pasteurella haemolytica A1 (strain B122).
XX
PN   CA2099707-A.
XX
PD   03-JAN-1994.
XX
PF   29-JUN-1993;   93CA-2099707.
XX
PR   02-JUL-1992;   92US-0908253.
XX
PA   (UYSA-) UNIV SASKATCHEWAN.
XX
PI   Harland RJ,  Potter AA;
XX
DR   WPI; 1994-092909/12.
DR   N-PSDB; AAQ44760.
XX
PT   Haemophilus somnus outer membrane protein extract -
PT   enriched with iron-regulated proteins, opt. contg.
PT   leuco:toxin antigens, for use as vaccine
XX
PS   Claim 5; Fig 5; 78pp; English.
XX
CC   A vaccine comprising an outer membrane protein (OMP) extract of
CC   Haemophilus somnus enriched with iron regulated proteins is new.
CC   The vaccine pref. further comprises an immunogenic leukotoxin
CC   polypeptide, esp. an immunogenic Pasteurella haemolytica leukotoxin
CC   homologous to LKT352.  Example 1.2 describes the prodn. of
CC   P. haemolytica recombinant leukotoxin from pAA352.
CC   Two expression constructs were made.  One, pAA342, contained the
CC   5'-AhaII fragment of the ltkA gene, while the other, pAA345,
CC   contained the entire lktA gene.  pAA342 expressed a truncated
CC   leukotoxin peptide at high levels, while pAA345 expressed full
CC   length leukotoxin at very low levels.  Therefore, the 3' end
CC   of the lktA gene was ligated into pAA342, yielding plasmid pAA352.
CC   LKT352 or new leukotoxin is 98% homologous to authentic
CC   leukotoxin.
CC   NB: the protein sequence in Fig 5 comprises 926 amino acids,
CC   however this protein is described in the text as having
CC   931 amino acids.
XX
SQ   Sequence   926 AA;

  Query Match             1.1%;  Score 10;  DB 15;  Length 926;
  Best Local Similarity   100.0%;  Pred. No. 3.7;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        265 LAQRVAAGLS 274
              ||||||||||
Db        256 LAQRVAAGLS 265


RESULT 28
AAW03945
ID   AAW03945 standard; Protein; 926 AA.
XX
AC   AAW03945;
XX
DT   20-NOV-1996  (first entry)
XX
DE   P. haemolytica truncated leukotoxin (LKT352).
XX
KW   Leukotoxin; LKT; gonadotropin-releasing hormone; GnRH;
KW   fusion protein; immunogen; vaccine; fertility control;
KW   contraceptive; sterilisation; plasmid pAA352.
XX
OS   Pasteurella haemolytica A1 strain B122.
XX
PN   WO9624675-A1.
XX
PD   15-AUG-1996.
XX
PF   24-JAN-1996;   96WO-CA00049.
XX
PR   10-FEB-1995;   95US-0387156.
XX
PA   (UYSA-) UNIV SASKATCHEWAN.
XX
PI   Manns JG,  Potter AA;
XX
DR   WPI; 1996-384447/38.
DR   N-PSDB; AAT37179.
XX
PT   Gonadotropin-releasing hormone multimer fusion proteins - with
PT   leukotoxin polypeptide for increased immunogenicity, useful in
PT   antifertility vaccine prodn.
XX
PS   Example 1; Fig 3A-3I; 87pp; English.
XX
CC   A truncated leukotoxin (AAW03945), LKT 352, lacks the cytotoxic
CC   portion of the native protein from Pasteurella haemolytica.  It is
CC   the product of plasmid pAA352 which carries a truncated lktA gene
CC   (AAT37179).  A fusion protein (AAW03942) between LKT352 and a
CC   gonadotropin releasing hormone tetramer can be expressed in
CC   Escherichia coli.  This is useful as a vaccine for fertility
CC   control, partic. immunological sterilisation of domestic or
CC   farm animals.
XX
SQ   Sequence   926 AA;

  Query Match             1.1%;  Score 10;  DB 17;  Length 926;
  Best Local Similarity   100.0%;  Pred. No. 3.7;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        265 LAQRVAAGLS 274
              ||||||||||
```

```
Db        256 LAQRVAAGLS 265


RESULT 29
AAW79568
ID   AAW79568 standard; Protein; 926 AA.
XX
AC   AAW79568;
XX
DT   24-DEC-1998  (first entry)
XX
DE   Leukotoxin 352 polypeptide.
XX
KW   Gonadotropin releasing hormone; GnHR; chimera; leukotoxin polypeptide;
KW   multimer; vaccine; tumour; Leukotoxin 352; LKT 352; lktA; plasmid pAA352;
KW   cytotoxic.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Misc_feature    11..924
FT                   /note= "Recombinant leukotoxin peptide"
XX
PN   WO9806848-A1.
XX
PD   19-FEB-1998.
XX
PF   08-AUG-1997;   97WO-CA00559.
XX
PR   09-AUG-1996;   96US-0694865.
XX
PA   (UYSA-) UNIV SASKATCHEWAN.
XX
PI   Manns JG,  Potter AA;
XX
DR   WPI; 1998-159540/14.
DR   N-PSDB; AAV61530.
XX
PT   Chimeric protein of leukotoxin and gonadotropin releasing hormone  -
PT   useful for, e.g. preparation of vaccines for reduction of incidence
PT   of mammary tumours in mammals
XX
PS   Claim 7; Figure 3.1-9; 118pp; English.
XX
CC   The present sequence represents a recombinantly produced or chemically
CC   synthesised leukotoxin 352 (LKT 352)polypeptide, derived from the lktA
CC   gene that is present in the plasmid pAA352.  This gene produces a
CC   truncated protein that has an estimated molecular weight of about 99 kDa
CC   and lacks the cytotoxic portion of the molecule.  Thus this gene has a
CC   higher expression level than that of the full-length molecule.  This can
CC   be used in the construction of a chimeric protein that comprises a
CC   leukotoxin polypeptide, several multimers, and a GnRH sequence.  The
CC   chimeric protein can be used as a vaccine to help reduce the incidence of
CC   mammary tumours in a mammalian individual.
XX
SQ   Sequence   926 AA;

  Query Match            1.1%;  Score 10;  DB 19;  Length 926;
  Best Local Similarity   100.0%;  Pred. No. 3.7;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        265 LAQRVAAGLS 274
              ||||||||||
Db        256 LAQRVAAGLS 265


RESULT 30
AAW07637
ID   AAW07637 standard; Protein; 934 AA.
XX
AC   AAW07637;
XX
DT   25-MAR-2003  (updated)
DT   05-FEB-1997  (first entry)
XX
DE   P. suis leukotoxin gene pslktA, protein product.
XX
KW   Leukotoxin; pslktC; pslktA; pslktB; calcium dependent; RTX;
KW   cytotoxic protein; repeat of toxin; BL-3 cell; pig; lymphocyte;
KW   immunogen; vaccine; immunisation; diarrhoea; reagent;
KW   diagnosis; assay; enteritis.
XX
OS   Pasteurella suis.
XX
FH   Key             Location/Qualifiers
FT   Domain          229..245
FT                   /label= transmembrane_domain
FT   Domain          295..311
FT                   /label= transmembrane_domain
FT   Domain          377..393
FT                   /label= transmembrane_domain
FT   Region          710..718
FT                   /note= "glycine rich repeat sequence"
FT   Region          719..727
FT                   /note= "glycine rich repeat sequence"
FT   Region          728..736
FT                   /note= "glycine rich repeat sequence"
FT   Region          737..745
FT                   /note= "glycine rich repeat sequence"
FT   Region          746..754
FT                   /note= "glycine rich repeat sequence"
FT   Region          756..763
FT                   /note= "glycine rich repeat sequence"
FT   Region          764..772
FT                   /note= "glycine rich repeat sequence"
FT   Region          773..781
FT                   /note= "glycine rich repeat sequence"
XX
PN   US5559008-A.
XX
PD   24-SEP-1996.
XX
PF   22-MAR-1994;   94US-0215805.
XX
PR   05-NOV-1992;   92US-0972157.
PR   22-MAR-1994;   94US-0215805.
XX
PA   (CORR ) CORNELL RES FOUND INC.
XX
PI   Chang Y;
XX
DR   WPI; 1996-442408/44.
DR   N-PSDB; AAT45417.
XX
PT   Leukotoxin genes from Pasteurella suis - also vectors and cells for
PT   expressing gene prods., for use in vaccines against porcine
PT   enteritis
XX
PS   Claim 34; Columns 59-64; 47pp; English.
XX
CC   The present sequence is protein prod. of the Pasteurella suis
CC   leukotoxin gene pslktA, which together with the pslktC, and pslktB
CC   gene prods., comprises a high mol. wt. calcium dependent cytotoxic
CC   protein of the repeat of toxin family (RTX), capable of killing
CC   both BL-3 cells and pig lymphocytes. The leukotoxin proteins, and
CC   the pslktC, A and B genes, or fragments, can be used as immunogens
CC   in vaccines for the immunisation of pigs against diarrhoea, or as
CC   reagents in diagnostic assays. The leukotoxin genes were isolated
CC   by screening a P. suis genomic library, in the phage vector
CC   lambda-dash, with a DNA probe derived from pYFC19, a plasmid
CC   carrying the IKtCA locus (Chang et al., Infect. Immun., vol. 55,
CC   pp2348-2354, 1987).
CC   (Updated on 25-MAR-2003 to correct PF field.)
XX
SQ   Sequence   934 AA;
```

```
  Query Match             1.1%;  Score 10;  DB 17;  Length 934;
  Best Local Similarity   100.0%;  Pred. No. 3.8;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        265 LAQRVAAGLS 274
              ||||||||||
Db        279 LAQRVAAGLS 288


RESULT 31
AAR34547
ID   AAR34547 standard; Protein; 936 AA.
XX
AC   AAR34547;
XX
DT   25-MAR-2003  (updated)
DT   23-AUG-1993  (first entry)
XX
DE   GnRH-leukotoxin gene fusion prod.
XX
KW   Vector; LKT 352; flanking; recombinant; antigen; somatostatin;
KW   gonadotropin releasing hormone; rotavirus viral protein 4;
KW   carrier protein; lactation; reproduction; SRIF.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Protein         1..926
FT                   /note= "recombinant leukotoxin protein"
FT   Peptide         927..936
FT                   /note= "GnRH"
XX
PN   WO9308290-A1.
XX
PD   29-APR-1993.
XX
PF   15-OCT-1992;   92WO-CA00449.
XX
PR   16-OCT-1991;   91US-0779171.
PR   14-OCT-1992;   92US-0960932.
XX
PA   (UYSA-) UNIV SASKATCHEWAN.
XX
PI   Hughes HPA,  Potter AA,  Redmond MJ;
XX
DR   WPI; 1993-152482/18.
DR   N-PSDB; AAQ41322.
XX
PT   Immunological carrier system with enhanced immunogenicity -
PT   comprises chimeric protein comprising leuco:toxin peptide or
PT   homologous protein fused to antigen esp. somatostatin or
PT   gonadotropin releasing hormone
XX
PS   Example 2; Fig 8; 95pp; English.
XX
CC   Oligonucleotides contg. sequences from bovine gonadotrophin
CC   releasing hormone (GnRH) gene were constructed on a Pharmacia Gene
CC   Assembler using standard phosporamidite chemistry.  The oligo-
CC   nucleotides were annealed and ligated into vector pAA352 (contg.
CC   the Pateurella heamolytica leuko-toxin gene) which had been digested
CC   with BamHI.  The ligated DNA was used to transform E. coli strain
CC   MH3000.  Transformants contg. the oligonucleotide inserts were
CC   identified by restiction endonuclease mapping and the recombinant
CC   plasmid designated pAA502.  The chimeric protein produced from the
CC   plasmid works to bring about a larger immune response than the antigen
CC   alone, i.e. the leukotoxin works as a carrier protein.
CC   See also AAR34545-8.
CC   (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ   Sequence   936 AA;

  Query Match             1.1%;  Score 10;  DB 14;  Length 936;
  Best Local Similarity   100.0%;  Pred. No. 3.8;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        265 LAQRVAAGLS 274
              ||||||||||
Db        256 LAQRVAAGLS 265


RESULT 32
AAR34546
ID   AAR34546 standard; Protein; 943 AA.
XX
AC   AAR34546;
XX
DT   25-MAR-2003  (updated)
DT   23-AUG-1993  (first entry)
XX
DE   Somatostatin-leukotoxin gene fusion prod.
XX
KW   Vector; LKT 352; flanking; recombinant; antigen; somatostatin;
KW   gonadotropin releasing hormone; rotavirus viral protein 4;
KW   carrier protein; lactation; reproduction; SRIF.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Protein         1..929
FT                   /note= "recombinant leukotoxin protein"
FT   Peptide         930..943
FT                   /note= "SRIF"
XX
PN   WO9308290-A1.
XX
PD   29-APR-1993.
XX
PF   15-OCT-1992;   92WO-CA00449.
XX
PR   16-OCT-1991;   91US-0779171.
PR   14-OCT-1992;   92US-0960932.
XX
PA   (UYSA-) UNIV SASKATCHEWAN.
XX
PI   Hughes HPA,  Potter AA,  Redmond MJ;
XX
DR   WPI; 1993-152482/18.
DR   N-PSDB; AAQ41321.
XX
PT   Immunological carrier system with enhanced immunogenicity -
PT   comprises chimeric protein comprising leuco:toxin peptide or
PT   homologous protein fused to antigen esp. somatostatin or
PT   gonadotropin releasing hormone
XX
PS   Example 2; Fig 6; 95pp; English.
XX
CC   Oligonucleotides contg. sequences from bovine somatostatin (SRIF)
CC   gene were constructed on a Pharmacia Gene Assembler using standard
CC   phosporamidite chemistry.  The oligonucleotides were annealed and
CC   ligated into vector pAA352 (contg. the Pateurella heamolytica leuko-
CC   toxin gene) which had been digested with BamHI.  The ligated DNA was
CC   used to transform E. coli strain JM105.  Transformants contg. the
CC   oligonucleotide inserts were identified by restiction endonuclease
CC   mapping and the recombinant plasmid designated pAA496.  The
CC   chimeric protein produced from the plasmid works to bring about a
CC   larger immune response than the antigen alone, i.e. the leukotoxin
CC   works as a carrier protein.
CC   See also AAR34545-8.
CC   (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ   Sequence   943 AA;

  Query Match             1.1%;  Score 10;  DB 14;  Length 943;
  Best Local Similarity   100.0%;  Pred. No. 3.8;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy        265 LAQRVAAGLS 274
              ||||||||||
Db        256 LAQRVAAGLS 265


RESULT 33
AAR34548
ID   AAR34548 standard; Protein; 951 AA.
XX
AC   AAR34548;
XX
DT   25-MAR-2003  (updated)
DT   23-AUG-1993  (first entry)
XX
DE   Rotavirus VP4-leukotoxin gene fusion prod.
XX
KW   Vector; LKT 352; flanking; recombinant; antigen; somatostatin;
KW   gonadotropin releasing hormone; rotavirus viral protein 4;
KW   carrier protein; lactation; reproduction; SRIF.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Protein         1..926
FT                   /note= "recombinant leukotoxin protein"
FT   Peptide         927..951
FT                   /note= "Rotavirus VP-4"
XX
PN   WO9308290-A1.
XX
PD   29-APR-1993.
XX
PF   15-OCT-1992;   92WO-CA00449.
XX
PR   16-OCT-1991;   91US-0779171.
PR   14-OCT-1992;   92US-0960932.
XX
PA   (UYSA-) UNIV SASKATCHEWAN.
XX
PI   Hughes HPA,  Potter AA,  Redmond MJ;
XX
DR   WPI; 1993-152482/18.
DR   N-PSDB; AAQ41322.
XX
PT   Immunological carrier system with enhanced immunogenicity -
PT   comprises chimeric protein comprising leuco:toxin peptide or
PT   homologous protein fused to antigen esp. somatostatin or
PT   gonadotropin releasing hormone
XX
PS   Example 2; Fig 8; 95pp; English.
XX
CC   Oligonucleotides contg. sequences from bovine Rotavirus viral
CC   protein 4 (VP-4) gene were constructed on a Pharmacia Gene
CC   Assembler using standard phosporamidite chemistry.  The oligo-
CC   nucleotides were annealed and ligated into vector pAA352 (contg.
CC   the Pateurella heamolytica leuko-toxin gene) which had been digested
CC   with BamHI.  The ligated DNA was used to transform E. coli strain
CC   MH3000.  Transformants contg. the oligonucleotide inserts were
CC   identified by resticion endonuclease mapping and the recombinant
CC   plasmid designated pAA501.  The chimeric protein produced from the
CC   plasmid works to bring about a larger immune response than the antigen
CC   alone, i.e. the leukotoxin works as a carrier protein.
CC   See also AAR34545-7.
CC   (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ   Sequence   951 AA;

  Query Match             1.1%;  Score 10;  DB 14;  Length 951;
  Best Local Similarity   100.0%;  Pred. No. 3.8;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        265 LAQRVAAGLS 274
              ||||||||||
Db        256 LAQRVAAGLS 265

RESULT 34
AAR07167
ID   AAR07167 standard; protein; 953 AA.
XX
AC   AAR07167;
XX
DT   24-JAN-1991  (first entry)
XX
DE   105kD PTX protein of P.haemolytica.
XX
KW   Vaccine; shipping fever.
XX
OS   Pasteurella haemolytica.
XX
PN   US4957739-A.
XX
PD   18-SEP-1990.
XX
PF   13-AUG-1987;   87US-0085430.
XX
PR   13-AUG-1987;   87US-0085430.
XX
PA   (TEXA ) UNIV OF TEXAS SYSTE.
XX
PI   Berget P,  Engler M,  Highlander S,  Weinstock G;
XX
DR   WPI; 1990-304558/40.
DR   N-PSDB; AAQ06074.
XX
PT   Vaccine against shipping fever in cattle - contains purified
PT   Pasteurella haemolytica antigen of molecular wt. 105 kD
XX
PS   Disclosure; Fig 9a-h; 35pp; English.
XX
CC   Gene product is antigenic to P.haemolytica, and may be used as a
CC   vaccine for immunisation against shipping fever. Abs raised to the
CC   antigen may be used in passive immunisation and diagnosis.
XX
SQ   Sequence   953 AA;

  Query Match             1.1%;  Score 10;  DB 11;  Length 953;
  Best Local Similarity   100.0%;  Pred. No. 3.8;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        265 LAQRVAAGLS 274
              ||||||||||
Db        283 LAQRVAAGLS 292


RESULT 35
AAR15159
ID   AAR15159 standard; Protein; 953 AA.
XX
AC   AAR15159;
XX
DT   10-JAN-1992  (first entry)
XX
DE   Leukotoxin from P. haemolytica.
XX
KW   Leukotoxin; probe; pasteurellosis; cytotoxicity; leukocytes.
XX
OS   Pasteurella haemolytica A1 (ATCC 43270).
XX
PN   US5055400-A.
XX
PD   08-OCT-1991.
XX
PF   26-NOV-1986;   86US-0935493.
XX
PR   26-NOV-1986;   86US-0935493.
```

```
XX
PA   (UYGU-) UNIV OF GUELPH.
XX
PI   Lo RYC,  Shewen PE,  Strathdee CA;
XX
DR   WPI; 1991-317648/43.
DR   N-PSDB; AAQ14000.
XX
PT   DNA encoding Pasteurella haemolytica-A1 leukotoxin - used to
PT   produce recombinant protein for prepn. of vaccine and antibodies
PT   for protection of cattle from pneumonic pasteurellosis
XX
PS   Claim 6; Page 15; 20pp; English.
XX
CC   ORF2 (AAR14223) codes for a 19.820 D protein of 166 amino acids;
CC   ORF1 codes for leukotoxin, a 101.883 D protein of 952 amino acids,
CC   which is produced on expression of pLKT5 in a host organism.
CC   The leukotoxin has cytotoxic activity specific against leukocytes.
CC   The protein is used to prepare vaccines or to raise antibodies
CC   against P. haemolytica to protect cattle from pneumonic pasteurellosis.
XX
SQ   Sequence   953 AA;

  Query Match             1.1%;  Score 10;  DB 12;  Length 953;
  Best Local Similarity   100.0%;  Pred. No. 3.8;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        265 LAQRVAAGLS 274
              ||||||||||
Db        283 LAQRVAAGLS 292


RESULT 36
AAR43865
ID   AAR43865 standard; Protein; 953 AA.
XX
AC   AAR43865;
XX
DT   25-MAR-2003  (updated)
DT   22-DEC-1993  (first entry)
XX
DE   Leukotoxin protein.
XX
KW   Antigenic determinants; peptides; vaccine; supernatant; leukotoxin.
XX
OS   Pasteurella haemolytica.
XX
PN   CA2081950-A.
XX
PD   02-MAY-1993.
XX
PF   02-NOV-1992;   92CA-2081950.
XX
PR   01-NOV-1991;   91US-0786662.
XX
PA   (UYGU-) UNIV GUELPH.
XX
PI   Conlon JA,  Lo RYC,  Shewen PE,  Strathdee CA;
XX
DR   WPI; 1993-227822/29.
DR   N-PSDB; AAQ46410.
XX
PT   Vaccine for preventing Pasteurella haemolytica infections -
PT   contains leukotoxin protein antigenic determinants and P.
PT   haemolytica culture supernatant
XX
PS   Claim 1; Figure 1; 40pp; English.
XX
CC   One or more biologically pure antigenic determinants of the
CC   leukotoxin protein comprising of at least six amino acids is a
CC   component of a new vaccine for animals.  The vaccine elicits an
CC   enhanced immune response after challenge with Pasteurella
CC   haemolytica.  The vaccines other component is a bacterial free
CC   culture supernatant derived from a culture of Pasteurella
CC   haemolytica.
CC   (Updated on 25-MAR-2003 to correct PF field.)
XX
SQ   Sequence   953 AA;

  Query Match             1.1%;  Score 10;  DB 14;  Length 953;
  Best Local Similarity   100.0%;  Pred. No. 3.8;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        265 LAQRVAAGLS 274
              ||||||||||
Db        283 LAQRVAAGLS 292


RESULT 37
AAR60072
ID   AAR60072 standard; Protein; 953 AA.
XX
AC   AAR60072;
XX
DT   25-MAR-2003  (updated)
DT   13-MAR-1995  (first entry)
XX
DE   PtxA protein of Pasteurella haemolytica.
XX
KW   Pasteurella haemolytica; shipping fever; pasteurellosis; cattle;
KW   vaccine; treatment; therapy; prophylaxis; antigen; antibody;
KW   immunoassay.
XX
OS   Pasteurella haemolytica.
XX
PN   US5336491-A.
XX
PD   09-AUG-1994.
XX
PF   15-JUN-1992;   92US-0899100.
XX
PR   28-NOV-1986;   86US-0935806.
PR   13-AUG-1987;   87US-0085430.
PR   18-JUN-1990;   90US-0540261.
PR   15-JUN-1992;   92US-0899100.
XX
PA   (TEXA ) UNIV TEXAS SYSTEM.
XX
PI   Berget P,  Engler M,  Highlander S,  Weinstock G;
XX
DR   WPI; 1994-255245/31.
DR   N-PSDB; AAQ70050.
XX
PT   New purified antigen from Pasteurella haemolytica - useful in
PT   vaccines against shipping fever of cattle, also for raising
PT   antibodies useful in diagnosis and passive immunisation
XX
PS   Disclosure; Figure 9; 35pp; English.
XX
CC   This protein is a 105 kD antigen which may be used in vaccines to
CC   protect cattle against shipping fever.  The antigen is also useful
CC   for raising antisera which can be used for passive immunisation (for
CC   treatment or temporary prophylaxis) and also to raise antibodies
CC   which can be used in immunoassay methods for the detection of
CC   Pasteurella haemolytica antigens in biological fluids.
CC   (Updated on 25-MAR-2003 to correct PF field.)
XX
SQ   Sequence   953 AA;

  Query Match             1.1%;  Score 10;  DB 15;  Length 953;
  Best Local Similarity   100.0%;  Pred. No. 3.8;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        265 LAQRVAAGLS 274
              ||||||||||
Db        283 LAQRVAAGLS 292
```

```
RESULT 38
AAE04638
ID   AAE04638 standard; Protein; 953 AA.
XX
AC   AAE04638;
XX
DT   10-SEP-2001  (first entry)
XX
DE   Pasteurella haemolytica leukotoxin protein.
XX
KW   Leukotoxin; respiratory disease; infection; therapy; immunostimulant;
KW   antibacterial; vaccine; transgenic plant; animal feed.
XX
OS   Pasteurella haemolytica.
XX
FH   Key             Location/Qualifiers
FT   Domain          230..250
FT                   /label= Hydrophobic_transmembrane_domain_1
FT   Domain          280..320
FT                   /label= Hydrophobic_transmembrane_domain_2
FT   Domain          360..400
FT                   /label= Hydrophobic_transmembrane_domain_3
XX
PN   WO200144289-A2.
XX
PD   21-JUN-2001.
XX
PF   15-DEC-2000; 2000WO-CA01498.
XX
PR   17-DEC-1999;   99US-0172148.
XX
PA   (UYGU-) UNIV GUELPH OFFICE.
XX
PI   Lo RYC,  Shewen PE,  Lee RWH,  Hodgins D,  Strommer JN;
XX
DR   WPI; 2001-408470/43.
XX
PT   Modified leukotoxin polypeptide is useful in a vaccine to prevent or
PT   treat Mannheimia (Pasteurella) infection (particularly M. haemolytica
PT   infection), and disease associated with a leukotoxin, e.g., respiratory
PT   disease -
XX
PS   Claim 6; Fig 13; 70pp; English.
XX
CC   The present sequence is Pasteurella (Mannheimia) haemolytica full
CC   length leukotoxin protein. The present invention relates to modified
CC   leukotoxin DNA sequences, wherein the modification comprises the
CC   removal of DNA sequence encoding the amino acids within the hydrophobic
CC   transmembrane domain of a full length leukotoxin protein. Modified
CC   leukotoxin sequences are used in vaccines to treat or prevent diseases
CC   associated with leukotoxin, e.g., respiratory disease, and Mannheimia
CC   infection (particularly M. haemolytica infection). In addition, the
CC   vaccine is used to prepare a medicament. Furthermore, the plant
CC   transformed with modified leukotoxin sequences is fed to an animal
CC   such as a ruminant, to prevent or treat respiratory diseases.
XX
SQ   Sequence   953 AA;

  Query Match             1.1%;  Score 10;  DB 22;  Length 953;
  Best Local Similarity   100.0%;  Pred. No. 3.8;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        265 LAQRVAAGLS 274
              ||||||||||
Db        283 LAQRVAAGLS 292


RESULT 39
AAW03942
ID   AAW03942 standard; Protein; 977 AA.
XX
AC   AAW03942;
XX
DT   20-NOV-1996  (first entry)
XX
DE   LKT-GnRH protein fusion from pCB113.
XX
KW   Leukotoxin; LKT; gonadotropin-releasing hormone; GnRH;
KW   fusion protein; immunogen; vaccine; fertility control;
KW   contraceptive; sterilisation.
XX
OS   Chimeric Pasteurella haemolytica A1 strain B122;
OS   Chimeric synthetic.
XX
FH   Key             Location/Qualifiers
FT   Domain          1..929
FT                   /label= LKT
FT   Domain          927..977
FT                   /label= GnRH_repeat_domain
XX
PN   WO9624675-A1.
XX
PD   15-AUG-1996.
XX
PF   24-JAN-1996;   96WO-CA00049.
XX
PR   10-FEB-1995;   95US-0387156.
XX
PA   (UYSA-) UNIV SASKATCHEWAN.
XX
PI   Manns JG,  Potter AA;
XX
DR   WPI; 1996-384447/38.
DR   N-PSDB; AAT37176.
XX
PT   Gonadotropin-releasing hormone multimer fusion proteins - with
PT   leukotoxin polypeptide for increased immunogenicity, useful in
PT   antifertility vaccine prodn.
XX
PS   Claim 7; Fig 5A-5H; 87pp; English.
XX
CC   A chimeric protein (AAW03942) is composed of a fusion between
CC   a truncated leukotoxin (LKT-352) from Pasteurella haemolytica (see
CC   also AAW03945) and a 4-copy gonadoliberin-releasing hormone (GnRH)
CC   repeat sequence (see also AAW03944).  It is the product of a
CC   chimeric gene (AAT37176) produced by ligating a synthetic sequence
CC   for the 4-copy GnRH into vector pAA352 (ATCC 68283), which carries
CC   the LKT-352 gene.  Recombinant plasmid pCB113 (LKT 352:4 copy
CC   GnRH, ATCC 69749) was obtd.  Escherichia coli transformants
CC   produced the chimeric protein, which is useful as a vaccine for
CC   fertility control, esp. immunological sterilisation of
CC   domestic or farm animals.
XX
SQ   Sequence   977 AA;

  Query Match             1.1%;  Score 10;  DB 17;  Length 977;
  Best Local Similarity   100.0%;  Pred. No. 3.9;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        265 LAQRVAAGLS 274
              ||||||||||
Db        256 LAQRVAAGLS 265


RESULT 40
AAW79569
ID   AAW79569 standard; Protein; 977 AA.
XX
AC   AAW79569;
XX
DT   24-DEC-1998  (first entry)
XX
DE   LKT-GnRH chimeric protein.
XX
```

```
KW   Chimera; pCB113; LKT 352; GnRH; Gonadotropin releasing hormone; multimer;
KW   cytotoxic activity; antigen presentation; immune response; vaccine;
KW   tumour.
XX
OS   Synthetic.
XX
PN   WO9806848-A1.
XX
PD   19-FEB-1998.
XX
PF   08-AUG-1997;   97WO-CA00559.
XX
PR   09-AUG-1996;   96US-0694865.
XX
PA   (UYSA-) UNIV SASKATCHEWAN.
XX
PI   Manns JG,  Potter AA;
XX
DR   WPI; 1998-159540/14.
DR   N-PSDB; AAV61531.
XX
PT   Chimeric protein of leukotoxin and gonadotropin releasing hormone  -
PT   useful for, e.g. preparation of vaccines for reduction of incidence
PT   of mammary tumours in mammals
XX
PS   Disclosure; Figure 5.1-8; 118pp; English.
XX
CC   The present sequence represents the LKT-GnRH chimeric protein from
CC   pCB113.  This plasmid contains the LKT 352 polypeptide (AAW79568) fused
CC   to four copies of the GnRH peptide.  This chimera lacks cytotoxic
CC   activity which enables there to be an increase in antigen presentation
CC   and thus an optimal immune response.  The removal of this region also
CC   enables the truncated LKT to be expressed at much higher levels and
CC   allows the amount of antigen administered to be reduced.  This chimeric
CC   protein comprises a leukotoxin polypeptide, several multimers, and a GnRH
CC   sequence.  The chimeric protein can be used as a vaccine to help reduce
CC   the incidence of mammary tumours in a mammalian individual.
XX
SQ   Sequence   977 AA;

  Query Match             1.1%;  Score 10;  DB 19;  Length 977;
  Best Local Similarity   100.0%;  Pred. No. 3.9;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        265 LAQRVAAGLS 274
              ||||||||||
Db        256 LAQRVAAGLS 265

Search completed: February 17, 2004, 10:19:35
Job time : 66 secs
```

OM protein - protein search, using sw model

Run on: February 17, 2004, 10:09:50 ; Search time 17 Seconds
(without alignments)
2564.339 Million cell updates/sec

Title: US-10-069-799-5
Perfect score: 4647
Sequence: 1 MSNINVIKSNIQAGLNSTKS..........SSNALQPITQPTQGILAPSV 927

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_41:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2334	50.2	953	1	LKA1_PASHA	P16535 pasteurella
2	2332	50.2	953	1	LKAB_PASHA	P55118 pasteurella
3	2306	49.6	947	1	LKTA_PASSP	P55123 pasteurella
4	2256	48.5	953	1	LKA3_PASHA	P55116 pasteurella
5	2236.5	48.1	956	1	RT2A_ACTPL	P15377 actinobacil
6	2236	48.1	955	1	LKAA_PASHA	P55117 pasteurella
7	2214.5	47.7	956	1	HLYA_ACTSU	Q00951 actinobacil
8	1955.5	42.1	1052	1	RT32_ACTPL	P55131 actinobacil
9	1924	41.4	1049	1	RT31_ACTPL	P55130 actinobacil
10	1848.5	39.8	1024	1	HLYA_ECOLI	P08715 escherichia
11	1845	39.7	1023	1	HLY1_ECOLI	P09983 escherichia
12	1729.5	37.2	1023	1	RT12_ACTPL	P55129 actinobacil
13	1723	37.1	1023	1	RT11_ACTPL	P55128 actinobacil
14	1712.5	36.9	1050	1	LKTA_ACTAC	P16462 actinobacil
15	802	17.3	1706	1	CYAA_BORPE	P15318 bordetella
16	801.5	17.2	1705	1	CYAA_BORBR	Q57506 bordetella
17	437	9.4	1829	1	FRPC_NEIMB	Q9jyv5 neisseria m
18	435	9.4	1829	1	FRPC_NEIMC	P55127 neisseria m
19	421	9.1	1302	1	FRPA_NEIMB	Q9k0k9 neisseria m
20	408.5	8.8	1115	1	FRPA_NEIMC	P55126 neisseria m
21	189.5	4.1	1608	1	HLYA_SERMA	P15320 serratia ma
22	188.5	4.1	2869	1	RBP1_PLAVB	Q00798 plasmodium
23	183.5	3.9	1577	1	HLYA_PROMI	P16466 proteus mir
24	180.5	3.9	1683	1	YL24_ANASP	Q8yv57 anabaena sp
25	178.5	3.8	284	1	NODO_RHILV	P15728 rhizobium l
26	178.5	3.8	1025	1	SLAP_CAUCR	P35828 caulobacter
27	175.5	3.8	1228	1	SLAP_BACST	P35825 bacillus st
28	173	3.7	1251	1	RBP2_PLAVB	Q00799 plasmodium
29	170	3.7	491	1	ZAPA_PROMI	Q11137 proteus mir
30	169.5	3.6	2249	1	OMPA_RICRI	P15921 rickettsia
31	169	3.6	1520	1	PMPD_CHLMU	Q9p1b0 chlamydia m
32	167	3.6	1656	1	OMPB_RICJA	O06653 r outer mem
33	165	3.6	479	1	PRTC_ERWCH	P16317 erwinia chr
34	164.5	3.5	475	1	PRTG_ERWCH	Q07162 erwinia chr
35	163.5	3.5	1592	1	GTF2_STRDO	P27470 streptococc
36	163	3.5	918	1	YMJB_CAEEL	P34487 caenorhabdi
37	163	3.5	1713	1	LMA3_HUMAN	Q16787 homo sapien
38	162.5	3.5	2003	1	YDBA_ECOLI	P33666 escherichia
39	161	3.5	2366	1	TOXB_CLODI	P18177 clostridium
40	160.5	3.5	1609	1	PM21_CHLPN	Q9z6u5 chlamydia p
41	160	3.4	1250	1	YFAL_ECOLI	P45508 escherichia
42	159.5	3.4	487	1	PRZN_SERMA	P23694 serratia ma
43	159	3.4	1569	1	YPJA_ECOLI	P52143 escherichia
44	157	3.4	1655	1	OMPB_RICCN	Q9kka3 r outer mem
45	156.5	3.4	2358	1	YEEJ_ECOLI	P76347 escherichia

ALIGNMENTS

```
RESULT 1
LKA1_PASHA
ID   LKA1_PASHA     STANDARD;      PRT;   953 AA.
AC   P16535;
DT   01-AUG-1990 (Rel. 15, Created)
DT   01-AUG-1990 (Rel. 15, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Leukotoxin from serotype A1.
GN   LKTA.
OS   Pasteurella haemolytica.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC   Pasteurellaceae; Mannheimia.
OX   NCBI_TaxID=75985;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Serotype A1;
RX   MEDLINE=87306837; PubMed=3040588;
RA   Lo R.Y.C., Strathdee C.A., Shewen P.E.;
RT   "Nucleotide sequence of the leukotoxin genes of Pasteurella
RT   haemolytica A1.";
RL   Infect. Immun. 55:1987-1996(1987).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Serotype A1 / PHL101;
RX   MEDLINE=89210283; PubMed=2707120;
RA   Highlander S.K., Chidambaram M., Engler M.J., Weinstock G.M.;
RT   "DNA sequence of the Pasteurella haemolytica leukotoxin gene
RT   cluster.";
RL   DNA 8:15-28(1989).
RN   [3]
RP   SEQUENCE OF 884-953 FROM N.A.
RC   STRAIN=Serotype A1 / PHL101;
RX   MEDLINE=90236888; PubMed=2185213;
RA   Highlander S.K., Engler M.J., Weinstock G.M.;
RT   "Secretion and expression of the Pasteurella haemolytica Leukotoxin.";
RL   J. Bacteriol. 172:2343-2350(1990).
CC   -!- FUNCTION: Bacterial hemolysins are exotoxins that attack blood
CC       cell membranes and cause cell rupture by mechanisms not clearly
CC       defined.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- DOMAIN: The Gly-rich region is probably involved in binding
CC       calcium, which is required for target cell-binding or cytolytic
CC       activity.
CC   -!- DOMAIN: The three transmembrane domains are believed to be
CC       involved in pore formation by the cytotoxin (By similarity).
CC   -!- PTM: Palmitoylated by lktC. The toxin only becomes active when
CC       modified (By similarity).
CC   -!- SIMILARITY: BELONGS TO THE RTX PROKARYOTIC TOXIN FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
```

```
CC   -----------------------------------------------------------------
DR   EMBL; M20730; AAA25529.1; -.
DR   EMBL; M24197; AAA25543.1; -.
DR   PIR; B30169; B30169.
DR   InterPro; IPR001343; Hemlysn_Ca_bind.
DR   InterPro; IPR003995; RtxA.
DR   Pfam; PF00353; hemolysinCabind; 5.
DR   Pfam; PF02382; RTX; 1.
DR   PRINTS; PR00313; CABNDNGRPT.
DR   PRINTS; PR01488; RTXTOXINA.
DR   PROSITE; PS00330; HEMOLYSIN_CALCIUM; 4.
KW   Hemolysis; Toxin; Cytolysis; Repeat; Calcium; Transmembrane;
KW   Lipoprotein; Palmitate.
FT   TRANSMEM    229    249       POTENTIAL.
FT   TRANSMEM    297    318       POTENTIAL.
FT   TRANSMEM    366    390       POTENTIAL.
FT   DOMAIN      734    784       6 X REPEATS, GLY-RICH (BY SIMILARITY).
FT   REPEAT      734    739       1.
FT   REPEAT      743    748       2.
FT   REPEAT      752    757       3.
FT   REPEAT      761    766       4.
FT   REPEAT      770    775       5.
FT   REPEAT      779    784       6.
FT   CONFLICT    409    414       FEHVAN -> LSTLQI (IN REF. 2).
FT   CONFLICT    742    742       D -> Y (IN REF. 2).
SQ   SEQUENCE   953 AA;  101996 MW;  7F93D113A118C05F CRC64;

  Query Match             50.2%;  Score 2334;  DB 1;  Length 953;
  Best Local Similarity   50.1%;  Pred. No. 5.6e-110;
  Matches  465;  Conservative  175;  Mismatches  263;  Indels   26;  Gaps   13;

Qy          8 KSNIQAGLNSTKSGLKNLYLAIPKD--YDPQKGGTLNDFIKAADELGIARLAEEPNHTET 65
Db         28 QSLTQAG-SSLKTGAKKIILYIPQNYQYDTEQGNGLQDLVKAAEELGIEVQREERNNIAT 86

Qy         66 AKKSVDTVNQFLSLTQTGIAISATKLEKFLQKHSTNKLAKGLDSVENIDRKLGKASNVLS 125
Db         87 AQTSLGTIQTAIGLTERGIVLSAPQIDKLLQK---TKAGQALGSAESIVQNANKAKTVLS 143

Qy        126 TLSSFLGTALAGIELDSLIKKGDAAPDALAKASIDLINEIIGNLSQSTQTIEAFSSQLAK 185
Db        144 GIQSILGSVLAGMDLDEAL-QNNSNQHALAKAGLELTNSLIENIANSVKTLDEFGEQISQ 202

Qy        186 LGSTISQAKGFSNIGNKLQNL-NFSKTNLGLEIITGLLSGISAGFALADKNASTGKKVAA 244
Db        203 FGSKLQNIKGLGTLGDKLKNIGGLDKAGLGLDVISGLLSGATAALVLADKNASTAKKVGA 262

Qy        245 GFELSNQVIGNVTKAISSYVLAQRVAAGLSTTGAVAALITSSIMLAISPLAFMNAADKFN 304
Db        263 GFELANQVVGNITKAVSSYILAQRVAAGLSSTGPVAALIASTVSLAISPLAFAGIADKFN 322

Qy        305 HANALDEFAKQFRKFGYDGDHLLAEYQRGVGTIEASLTTISTALGAVSAGVSAAAVGSAV 364
Db        323 HAKSLESYAERFKKLGYDGDNLLAEYQRGTGTIDASVTAINTALAAIAGGVSAAAAGSVI 382

Qy        365 GTPIALLVAGVTGLISGILEASKQAMFESVANRLQGKILEWEKQNGGQNYFDKGYDSRYA 424
Db        383 ASPIALLVSGITGVISTILQYSKQAMFEHVANKIHNKIVEWEKNNHGKNYFENGYDARYL 442

Qy        425 AYLANNLKFLSELNKELEAERVIAITQQRWDNNIGELAGITKLGERIKSGKAYADAFEDG 484
Db        443 ANLQDNMKFLLNLNKELQAERVIAITQQQWDNNIGDLAGISRLGEKVLSGKAYVDAFEEG 502

Qy        485 KKVEAGSNITLDAKTGIIDISNSNGKKTQALHFTSPLLTAGTESRERLTNGKYSYINKLK 544
Db        503 KHIKADKLVQLDSANGIIDVSNSGKAKTQHILFRTPLLTPGTEHRERVQTGKYEYITKLN 562

Qy        545 FGRVKNWQVTDGEASSKLDFSKVIQRV-------AETEGTDEIGLIVNAKAGNDDIFVGQ 597
Db        563 INRVDSWKITDGAASSTFDLTNVVQRIGIELDNAGNVTKTKETKIIAKLGEGDDNVFVGS 622

Qy        598 GKMNIDGGDGHDRVFYSKDGGFGNITVDGTSATEAGSYTVNRKVARGDIYHEVVKRQETK 657
Db        623 GTTEIDGGEGYDRVHYSR-GNYGALTIDATKETEQGSYTVNRFVETGKALHEVTSTHTAL 681

Qy        658 VGKRTETIQYRDYELRKVGYGYQSTDNLKSVEEVIGSQFNDVFKGSKFNDIFHSGEGDDL 717
Db        682 VGNREEKIEYR-HSNNQHHAGYYTKDTLKAVEEIIGTSHNDIFKGSKFNDAFNGGDGVDT 740

Qy        718 LDGGAGDDRLFGGKGNDRLSGDEGDDLLDGGSGDDVLNGGAGNDVYIFRKGDGNDTLYDG 777
Db        741 IDGNDGNDRLFGGKGDDILDGGNGDDFIDGGKGNDLLHGGKGDDIFVHRKGDGNDIITDS 800

Qy        778 TGNDKLAFADANISDIMIERTKEGIIVKRNDHSGSINIPRWY----ITSNLQNYQSNKTD 833
Db        801 DGNDKLSFSDSNLKDLTFEKVKHNLVI-TNSKKEKVTIQNWFREADFAKEVPNYKATK-D 858

Qy        834 HKIEQLIGKDGSYITSDQIDKILQDKKDGTVITSQELKKLADENKSQKLSASDIASSLNK 893
Db        859 EKIEEIIGQNGERITSKQVDDLI--AKGNGKITQDELSKVVDNYELLKHS-KNVTNSLDK 915

Qy        894 LVGSMALFGTANSVSSNALQPITQPTQGI 922
Db        916 LISSVSAFTSSNDSRNVLVAPTSMLDQSL 944

RESULT 2
LKAB_PASHA
ID   LKAB_PASHA     STANDARD;      PRT;   953 AA.
AC   P55118;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Leukotoxin from serotype A11.
GN   LKTA.
OS   Pasteurella haemolytica.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC   Pasteurellaceae; Mannheimia.
OX   NCBI_TaxID=75985;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Serotype A11;
RX   MEDLINE=94041617; PubMed=8225575;
RA   Burrows L.L., Olah-Winfield E., Lo R.Y.C.;
RT   "Molecular analysis of the leukotoxin determinants from Pasteurella
RT   haemolytica serotypes 1 to 16.";
RL   Infect. Immun. 61:5001-5007(1993).
CC   -!- FUNCTION: Bacterial hemolysins are exotoxins that attack blood
CC       cell membranes and cause cell rupture by mechanisms not clearly
CC       DEFINED.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- DOMAIN: The Gly-rich region is probably involved in binding
CC       calcium, which is required for target cell-binding or cytolytic
CC       activity.
CC   -!- DOMAIN: The three transmembrane domains are believed to be
CC       involved in pore formation by the cytotoxin (BY SIMILARITY).
CC   -!- PTM: Palmitoylated by lktC. The toxin only becomes active when
CC       modified (By similarity).
CC   -!- SIMILARITY: BELONGS TO THE RTX PROKARYOTIC TOXIN FAMILY.
CC   -----------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----------------------------------------------------------------
DR   EMBL; U01215; AAB36689.1; -.
DR   InterPro; IPR001343; Hemlysn_Ca_bind.
DR   InterPro; IPR003995; RtxA.
DR   Pfam; PF00353; hemolysinCabind; 5.
DR   Pfam; PF02382; RTX; 1.
DR   PRINTS; PR00313; CABNDNGRPT.
DR   PRINTS; PR01488; RTXTOXINA.
DR   PROSITE; PS00330; HEMOLYSIN_CALCIUM; 4.
```

```
KW   Hemolysis; Toxin; Cytolysis; Repeat; Calcium; Transmembrane;
KW   Lipoprotein; Palmitate.
FT   TRANSMEM    230    250       POTENTIAL.
FT   TRANSMEM    297    317       POTENTIAL.
FT   TRANSMEM    381    401       POTENTIAL.
FT   DOMAIN      734    784       6 X REPEATS, GLY-RICH.
FT   REPEAT      734    739       1.
FT   REPEAT      743    748       2.
FT   REPEAT      752    757       3.
FT   REPEAT      761    766       4.
FT   REPEAT      770    775       5.
FT   REPEAT      779    784       6.
SQ   SEQUENCE   953 AA;  102206 MW;  927FF56CFC884F12 CRC64;

  Query Match             50.2%;  Score 2332;  DB 1;  Length 953;
  Best Local Similarity   49.1%;  Pred. No. 7.1e-110;
  Matches  467;  Conservative  183;  Mismatches  261;  Indels   40;  Gaps   15;

Qy          1 MSNINV-IKSN---IQAGLNST-----------KSGLKNLYLAIPKD--YDPQKGGTLND 43
                ::||:  :||:    ::|||  |             |:| | : | ||||  || :|| |  |
Db          5 LTNISTNLKSSWLTAKSGLNRTGQSLAKAGQSLKTGAKKIILYIPKDYQYDTEKGNGLQD 64

Qy         44 FIKAADELGIARLAEEPNHTETAKKSVDTVNQFLSLTQTGIAISATKLEKFLQKHSTNKL 103
                :|||:||||     || |       |: |: |:    | ||: || :|| :|:| |||     |:
Db         65 LVKAAEELGIEVQKEEGNDIAKAQTSLGTIQNVLGLTERGIVLSAPQLDKLLQK---TKV 121

Qy        104 AKGLDSVENIDRKLGKASNVLSTLSSFLGTALAGIELDSLIKKGDAAPDALAKASIDLIN 163
                : : | ||: :      |  |||: |  ||: |||::||   ::| ::       |||| ::| |
Db        122 GQAIGSAENLTKGFSNAKTVLSGIQSILGSVLAGMDLDEALQK-NSNELTLAKAGLELTN 180

Qy        164 EIIGNLSQSTQTIEAFSSQLAKLGSTISQAKGFSNIGNKLQNLN-FSKTNLGLEIITGLL 222
                :| |:: |  :|::||  |: :||| :     || |::|:|: |: | ||:|||::::|||
Db        181 SLIENIANSVKTLDAFGDQINQLGSKLQNVKGLSSLGDKLKGLSGFDKTSLGLDVVSGLL 240

Qy        223 SGISAGFALADKNASTGKKVAAGFELSNQVIGNVTKAISSYVLAQRVAAGLSTTGAVAAL 282
                || :|      ||||||| :|| |||:|||:|:||:|:|||:|||||||||||:|| ||||
Db        241 SGATAALVLADKNASTSRKVGAGFELANQVVGNITKAVSSYILAQRVAAGLSSTGPVAAL 300

Qy        283 ITSSIMLAISPLAFMNAADKFNHANALDEFAKQFRKFGYDGDHLLAEYQRGVGTIEASLT 342
                | |:: |||||||   |||||| :|: :|::|:| |||||:|||||||| |||: |:|
Db        301 IASTVSLAISPLAFAGIADKFNHAKSLESYAERFKKLGYDGDNLLAEYQRGTGTIDRSVT 360

Qy        343 TISTALGAVSAGVSAAAVGSAVGTPIALLVAGVTGLISGILEASKQAMFESVANRLQGKI 402
                |:||| |:: |||||   || : :||||||:|:||:|| ||: ||||||| |||::  ||
Db        361 AINTALAAIAGGVSAAGRGSVIASPIALLVSGITGVISTILQYSKQAMFEHVANKIHNKI 420

Qy        403 LEWEKQNGGQNYFDKGYDSRYAAYLANNLKFLSELNKELEAERVIAITQQRWDNNIGELA 462
                :||||· | :|||: |||:|| | |  | :|:|||     |||||:|||||||||||:||||||:||
Db        421 VEWEKNNHGKNYFENGYDARYLANLQDNMKFLLNLNKELQAERVIAITQQQWDNNIGDLA 480

Qy        463 GITKLGERIKSGKAYADAFEDGKKVEAGSNITLDAKTGIIDISNSNGKKTQALHFTSPLL 522
                ||::|||:||:|||||:||||:|  :|  ||::|:||||:|||:|||   ||| : :|||
Db        481 GISRLGEKVLSGKAYVDAFEEGKHLKADKLVQLDSANGIIDVSNSGKAKTQDILFRTPLL 540

Qy        523 TAGTESRERLTNGKYSYINKLKFGRVKNWQVTDGEASSKLDFSKVIQRV-------AETE 575
                | ||: |||:  ||| || ||    || :|::||| |||   | :|:||:
Db        541 TPGTDDRERVQTGKYEYITKLNINRVDSWKITDGAASSTFDLTNVVQRIGIELDNAGNVT 600

Qy        576 GTDEIGLIVNAKAGNDDIFVGQGKMNIDGGDGHDRVFYSKDGGFGNITVDGTSATEAGSY 635
                | |   ::       ||:|::||| |      ||||:|:||| ||: | :| :|:| |  || |||
Db        601 KTKETKIVAKLGAGDDNVFVGSGTTEIDGGEGYDRVHYSR-GNYGALTIDATKETEQGSY 659

Qy        636 TVNRKVARGDIYHEVVKRQETKVGKRTETIQYRDYELRKVGYGYQSTDNLKSVEEVIGSQ 695
                |||| |   |    ||           || | |  |:|| :     :     || : | ||:|||:||:
Db        660 TVNRFVETGKALHEGTSTHTALVGNREEKIEYR-HSNNQHHAGYYTKDTLKAVEEIIGTS 718

Qy        696 FNDVFKGSKFNDIFHSGEGDDLLDGGAGDDRLFGGKGNDRLSGDEGDDLLDGGSGDDVLN 755
                ||:|||||||| |: |:| |  :||   |:||||||||:| :  |   |||  :||| |:|:|:
Db        719 HNDIFKGSKFNDAFNGGDGVDTIDGKDGNDRLFGGKGDDIIDGGNGDDFIDGGKGNDLLH 778

Qy        756 GGAGNDVYIFRKGDGNDTLYDGTGNDKLAFADANISDIMIERTKEGIIVKRNDHSGSINI 815
                || |:|:::: |:||||| : |    |||||:|:|:|: |:   |:  |   :::    |       :  |
Db        779 GGKGDDIFVHRQGDGNDIITDSDGNDKLSFSDSNLKDLTFEKVKHNLVI-TNSRKEKVTI 837

Qy        816 PRWY----ITSNLQNYQSNKTDHKIEQLIGKDGSYITSDQIDKILQDKKDGTVITSQELK 871
                 |:          ::||::  |  | |||::||::|  |||  |:| ::    |     ||  ||
Db        838 QDWFREADFAKEVRNYKATK-DEKIEEIIGQNGERITSKQVDDLI--AKGNGKITQDELS 894

Qy        872 KLADENKSQKLSASDIASSLNKLVGSMALFGTANSVSSNALQPITQPTQGI 922
                 |: |   :   | |   :: :||:||:  |  :  |  ::|     :   :  |  :       |  :
Db        895 KVVDNYELLKHS-KNVTNSLDKLISSASAFTSSNDSRNVLVAPTSMLDQSL 944


RESULT 3
LKTA_PASSP
ID   LKTA_PASSP     STANDARD;      PRT;   947 AA.
AC   P55123;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Leukotoxin.
GN   LKTA.
OS   Pasteurella haemolytica-like sp. (strain 5943B).
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC   Pasteurellaceae; Pasteurella.
OX   NCBI_TaxID=28165;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93239320; PubMed=8478098;
RA   Chang Y.-F., Ma D.-P., Shi J., Chengappa M.M.;
RT   "Molecular characterization of a leukotoxin gene from a Pasteurella
RT   haemolytica-like organism, encoding a new member of the RTX toxin
RT   family.";
RL   Infect. Immun. 61:2089-2095(1993).
CC   -!- FUNCTION: Virulence factor which is cytotoxic for leukocytes but
CC       is not hemolytic.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- DOMAIN: The Gly-rich region is probably involved in binding
CC       calcium, which is required for target cell-binding or cytolytic
CC       activity.
CC   -!- DOMAIN: The three transmembrane domains are believed to be
CC       involved in pore formation by the cytotoxin (By similarity).
CC   -!- PTM: Palmitoylated by lktC. The toxin only becomes active when
CC       modified (By similarity).
CC   -!- SIMILARITY: BELONGS TO THE RTX PROKARYOTIC TOXIN FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; L12148; AAA16444.1; -.
DR   InterPro; IPR001343; Hemlysn_Ca_bind.
DR   InterPro; IPR003995; RtxA.
DR   Pfam; PF00353; hemolysinCabind; 5.
DR   Pfam; PF02382; RTX; 1.
DR   PRINTS; PR00313; CABNDNGRPT.
DR   PRINTS; PR01488; RTXTOXINA.
DR   PROSITE; PS00330; HEMOLYSIN_CALCIUM; 4.
KW   Hemolysis; Toxin; Cytolysis; Repeat; Calcium; Transmembrane;
KW   Lipoprotein; Palmitate.
FT   TRANSMEM    154    170       POTENTIAL.
FT   TRANSMEM    312    333       POTENTIAL.
FT   TRANSMEM    393    414       POTENTIAL.
FT   DOMAIN      625    780       7 X REPEATS, GLY-RICH.
FT   REPEAT      625    630       1.
FT   REPEAT      730    735       2.
FT   REPEAT      739    744       3.
FT   REPEAT      748    753       4.
FT   REPEAT      757    762       5.
FT   REPEAT      766    771       6.
FT   REPEAT      775    780       7.
SQ   SEQUENCE   947 AA;  101559 MW;  9744F06395EF5BED CRC64;
```

```
  Query Match          49.6%;  Score 2306;  DB 1;  Length 947;
  Best Local Similarity   50.0%;  Pred. No. 1.4e-108;
  Matches  466;  Conservative  173;  Mismatches  253;  Indels   40;  Gaps   16;

Qy          1 MSNINV-IKSNIQAGLNST-----------KSGLKNLYLAIPKD--YDPQKGGTLNDFIK 46
              ::||:  :|:::|:||:  |             |:| | | | ||||  ||   :|  | | :|
Db          4 LANISTNLKNSLQSGLHKTGQSLNQAGQSLKAGAKKLILYIPKDYEYDSGRGNGLQDLVK 63

Qy         47 AADELGIARLAEEPNHTETAKKSVDTVNQFLSLTQTGIAISATKLEKFLQKHSTNKLAKG 106
              ||::|||      || |    ||: |: |:    |   :: |: :|| :|:| |||:     |::|
Db         64 AAEDLGIEVQREERNGIATAQNSLSTIQNILGFSERGVVLSAPQLDKLLQKY---KISKA 120

Qy        107 LDSVENIDRKLGKASNVLSTLSSFLGTALAGIELDSLIKKGDAAPDALAKASIDLINEII 166
                | ||: :.|| |  :|| : | ||: :||::|| ::|     :  | |||| ::| | :|
Db        121 PGSSENVAKNLGNAQTLLSGIQSILGSVMAGMDLDEILKNKGSELD-LAKAGLELTNSLI 179

Qy        167 GNLSQSTQTIEAFSSQLAKLGSTISQAKGFSNIGNKLQNLN-FSKTNLGLEIITGLLSGI 225
                |:: | ||:: || |:::||: :       ||      :|:||:| : |||  ||||:|:||||
Db        180 ENIANSVQTLDTFSEQISQLGTKLQNVKGLGTLGDKLKNFSGFSKAGLGLEVISGLLSGA 239

Qy        226 SAGFALADKNASTGKKVAAGFELSNQVIGNVTKAISSYVLAQRVAAGLSTTGAVAALITS 285
              :|   ||||||| :|| ||||:|||:|:|:|||:|||||||||| || |:||| |
Db        240 TAALVLADKNASTDRKVGAGFELANQVVGNITKAVSSYILAQRVAAGLSNTGPVSALIAS 299

Qy        286 SIMLAISPLAFMNAADKFNHANALDEFAKQFRKFGYDGDHLLAEYQRGVGTIEASLTTIS 345
              :: ||||||||   |||||:| ||: :|::|:| ||:|| |||||||| |||:||:| ::
Db        300 TVALAISPLAFAGIADKFNNAKALESYAERFKKLGYEGDSLLAEYQRGTGTIDASVTAVN 359

Qy        346 TALGAVSAGVSAAAVGSAVGTPIALLVAGVTGLISGILEASKQAMFESVANRLQGKILEW 405
              ||| |:| ||||||  ||  || |||||:|:|||:|| ||: |||||| ||.|::  ||::|
Db        360 TALAAISGGVSAAAAGSLVGAPIALLVSGITGIISTILQYSKQAMFEHVANKIHDKIVDW 419

Qy        406 EKQNGGQNYFDKGYDSRYAAYLANNLKFLSELNKELEAERVIAITQQRWDNNIGELAGIT 465
              ||:: |:|||: ||||||  |  | :|:: |   ||||:|||||  ||||:||||||  ||||:
Db        420 EKKHNGKNYFENGYDSRYLADLQDNMRQLQNLNKELQAERVIRITQQQWDNNIGNLAGIS 479

Qy        466 KLGERIKSGKAYADAFEDGKKVEAGSNITLDAKTGIIDISNSNGKKTQALHFTSPLLTAG 525
              :|||:: |||||||||||:||  ::|  |  |:  |:|:  | |  |||  : :||||  |
Db        480 RLGEKVMSGKAYADAFEEGKLIKADTFVQLDSATGVINTSKSDNVKTQHILFRTPLLTPG 539

Qy        526 TESRERLTNGKYSYINKLKFGRVKNWQVTDGEASSKLDFSKVIQRVA-------ETEGTD 578
               |:|||:  ||| || ||    ||  :|::|||  :|  |  : |:||:              |
Db        540 VENRERIQTGKYEYITKLNINRVDSWKITDGATNSTFDLTNVVQRIGIELDHADNVTKTK 599

Qy        579 EIGLIVNAKAGNDDIFVGQGKMNIDGGDGHDRVFYSKDGGFGNITVDGTSATEAGSYTVN 638
               |  :| |   ||||:|:| |   :|||:| ||| ||:  | :| :|:| |: :    |||||
Db        600 ETKIIANLGDGNDDVFIGSGTTEVDGGNGLDRVHYSR-GDYGALTIDATNESVQGSYTVK 658

Qy        639 RKVARGDIYHEVVKRQETKVGKRTETIQYRDYELRKVGYGYQSTDNLKSVEEVIGSQFND 698
              | |   |  |||    |    || | | |:|| :        || :|| ||||||:||:  ||
Db        659 RFVETGKALHEVTATQSVLVGSREEKIEYR-HSNNTQHAGYYTTDTLKSVEEIIGTSRND 717

Qy        699 VFKGSKFNDIFHSGEGDDLLDGGAGDDRLFGGKGNDRLSGDEGDDLLDGGSGDDVLNGGA 758
              :|||||:| || |:| | :|| ||:||||||||  | :|||  :||| |||:|:||
Db        718 IFKGSKFDDAFHGGDGVDNIDGNAGNDRLFGGKGFDIIDGGDGDDFIDGGQGDDILHGGK 777

Qy        759 GNDVYIFRKGDGNDTLYDGTGNDKLAFADANISDIMIERTKEGIIVKRNDHSGSINIPRW 818
              |||:   || |||:: |  |||:|:|||:|: |:  |:       :::  |     :  |  |
Db        778 GNDILCTVKG-GNDSISDSGGNDRLSFADSNLKDLTFEKVNHHLMI-TNVKKEKVTIQNW 835

Qy        819 Y----ITSNLQNYQSNKTDHKIEQLIGKDGSYITSDQIDKILQDKKDGTVITSQELKKLA 874
              :          : |||:  ||||::||: |  |||  |||::::  |    |    ||:::|
Db        836 FREADYAKTVHNYQAT-ADEKIEEIIGRQGERITSKQIDELIEKGKG--KIDQSELERIA 892

Qy        875 DENKSQKLSASDIAS-SLNKLVGSMALFGTAN 905
              :   |  |  |  || ||||||  |  |  ::|
Db        893 E--SSALLKESKFASNSLNKLVSSAGAFASSN 922


RESULT 4
LKA3_PASHA
ID   LKA3_PASHA     STANDARD;      PRT;   953 AA.
AC   P55116;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Leukotoxin from serotype T3.
GN   LKTA.
OS   Pasteurella haemolytica.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC   Pasteurellaceae; Mannheimia.
OX   NCBI_TaxID=75985;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Serotype T3;
RX   MEDLINE=94041617; PubMed=8225575;
RA   Burrows L.L., Lo R.Y.C., Olah-Winfield E.;
RT   "Molecular analysis of the leukotoxin determinants from Pasteurella
RT   haemolytica serotypes 1 to 16.";
RL   Infect. Immun. 61:5001-5007(1993).
CC   -!- FUNCTION: Bacterial hemolysins are exotoxins that attack blood
CC       cell membranes and cause cell rupture by mechanisms not clearly
CC       defined.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- DOMAIN: The Gly-rich region is probably involved in binding
CC       calcium, which is required for target cell-binding or cytolytic
CC       activity.
CC   -!- DOMAIN: The three transmembrane domains are believed to be
CC       involved in pore formation by the cytotoxin (BY SIMILARITY).
CC   -!- PTM: Palmitoylated by lktC. The toxin only becomes active when
CC       modified (By similarity).
CC   -!- SIMILARITY: BELONGS TO THE RTX PROKARYOTIC TOXIN FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; U01216; AAB36691.1; -.
DR   InterPro; IPR001343; Hemlysn_Ca_bind.
DR   InterPro; IPR003995; RtxA.
DR   Pfam; PF00353; hemolysinCabind; 5.
DR   Pfam; PF02382; RTX; 1.
DR   PRINTS; PR00313; CABNDNGRPT.
DR   PRINTS; PR01488; RTXTOXINA.
DR   PROSITE; PS00330; HEMOLYSIN_CALCIUM; 2.
KW   Hemolysis; Toxin; Cytolysis; Repeat; Calcium; Transmembrane;
KW   Lipoprotein; Palmitate.
FT   TRANSMEM    229    249       POTENTIAL.
FT   TRANSMEM    297    318       POTENTIAL.
FT   TRANSMEM    381    401       POTENTIAL.
FT   DOMAIN      734    784       6 X REPEATS, GLY-RICH.
FT   REPEAT      734    739       1.
FT   REPEAT      743    748       2.
FT   REPEAT      752    757       3.
FT   REPEAT      761    766       4.
FT   REPEAT      770    775       5.
FT   REPEAT      779    784       6.
SQ   SEQUENCE   953 AA;  101948 MW;  FDBDCE2FDC85FDF2 CRC64;

  Query Match          48.5%;  Score 2256;  DB 1;  Length 953;
  Best Local Similarity   48.7%;  Pred. No. 4.6e-106;
  Matches  463;  Conservative  174;  Mismatches  270;  Indels   44;  Gaps   17;

Qy          5 NVIKSNIQA---GLN-----------STKSGLKNLYLAIPKD--YDPQKGGTLNDFIKAA 48
               | |:| : |    |||              : |:|:| : | |||| ||    |  | | :|||
Db         10 NGIRSTLTATRGGLNRAGQSLTQAGQTLKNGVKKIILYIPKDYKYDSGSGNGLQDLVKAA 69

Qy         49 DELGIARLAEEPNHTETAKKSVDTVNQFLSLTQTGIAISATKLEKFLQKHSTNKLAKGLD 108
              :||||     || |     |: |: |:    || :|| :|| :|:| |||    ||: : |
Db         70 EELGIEVQKEEGNDIAKAQTSLGTIQNVLGLTERGIVLSAPQLDKLLQK---NKVGQALG 126
```

```
Qy     109 SVENIDRKLGKASNVLSTLSSFLGTALAGIELDSLIKKGDAAPDALAKASIDLINEIIGN 168
Db     127 SSESIAQNFSQAKTVLSGVQSILGSVLAGMDLDEAL-QNESDQLTLAKAGLELTNSLIEN 185

Qy     169 LSQSTQTIEAFSSQLAKLGSTISQAKGFSNIGNKLQNL-NFSKTNLGLEIITGLLSGISA 227
Db     186 IANSVQTLDAFSEQISQFGSKLQNVKGLGALGDKLKNIGGLDKAGLGLHVISGLLSGATA 245

Qy     228 GFALADKNASTGKKVAAGFELSNQVIGNVTKAISSYVLAQRVAAGLSTTGAVAALITSSI 287
Db     246 ALVLADKDASTAKKVGAGFELANQVVGNITKAVSSYILAQRVAARLSSTGPVAALIASTV 305

Qy     288 MLAISPLAFMNAADKFNHANALDEFAKQFRKFGYDGDHLLAEYQRGVGTIEASLTTISTA 347
Db     306 ALAISPLSFAGIADKFDRAKSLENYAERFKKLGYEGDSLLAEYQHGTGTIDASVTAINTA 365

Qy     348 LGAVSAGVSAAAVGSAVGTPIALLVAGVTGLISGILEASKQAMFESVANRLQGKILEWEK 407
Db     366 LAAIAGGVSAAAAGSVVASPIALLVSGITGVISTILQYSKQAMFEHVANKIHNKIVEWEK 425

Qy     408 QNGGQNYFDKGYDSRYAAYLANNLKFLSELNKELEAERVIAITQQRWDNNIGELAGITKL 467
Db     426 NNGGKNYFENGYDARYLANLQDNMKFLLNLNKELQAERVIAITQQQWDSNIGDLAGISRL 485

Qy     468 GERIKSGKAYADAFEDGKKVEAGSNITLDAKTGIIDISNSNGKKTQALHFTSPLLTAGTE 527
Db     486 GEKVLSGKAYVDAFEEGQHLKADKLVQLDSAKGIIDVSNTGEAKTQHILFRTPLLTPGTE 545

Qy     528 SRERLTNGKYSYINKLKFGRVKNWQVTDGEASSKLDFSKVIQRVA-------ETEGTDEI 580
Db     546 KRERVQTGKYEYITKLHINRVDSWQIKDGAASSTFDLTNVVQRIGVELDHAENVIKTKET 605

Qy     581 GLIVNAKAGNDDIFVGQGKMNIDGGDGHDRVFYSKDGGFGNITVDGTSATEAGSYTVNRK 640
Db     606 KIVATLGDGDDNVFVGSGTTEIDGGEGYDRVHYSR-GNYGALTIDATKEPEQGSYTVNRF 664

Qy     641 VARGDIYHEVVKRQETKVGKRTETIQYRDYELRKVGYGYQSTDNLKSVEEVIGSQFNDVF 700
Db     665 VESGKALHEVTSTHTALVGNREEKIEYR-HSNNQHHAGYYTKDTLKAVEEIIGTSHNDIF 723

Qy     701 KGSKFNDIFHSGEGDDLLDGGAGDDRLFGGKGNDRLSGDEGDDLLDGGSGDDVLNGGAGN 760
Db     724 KGSKFNDAFNGGDGVDTIDGNDGNDRLFGGKGDDIIDGGNGDDFIDGGKGNHLLHGGKGV 783

Qy     761 DVYIFRKGDGNDTLYDGTGNDKLAFADANISDIMIERTKEGIIVKRNDHSGSINIPRWY- 819
Db     784 DIFVHRQGDGNDSITESEGNDKLSFSDSNLKDLTFEKVNHHLVI-TNTKQEKVTIQNWFR 842

Qy     820 ---ITSNLQNYQSNKTDHKIEQLIGKDGSYITSDQIDKILQDKKDGTVITSQELKKLADE 876
Db     843 EAEFAKTIQNYVATR-DDKIEEIIGQNGERITSKQVDDLI--AKGNGKIAQSELTKVVDN 899

Qy     877 NKSQKLSASDIASSLNKLVGSMALFGTANSVSSNALQPITQPTQGILAPSV 927
Db     900 YQLLKYS-RDASNSLDKLISSASAFTSSND-SRNVL---ASPT-SMLDPSL 944

RESULT 5
RT2A_ACTPL
ID   RT2A_ACTPL     STANDARD;      PRT;    956 AA.
AC   P15377;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-APR-1990 (Rel. 14, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   RTX-II toxin determinant A (APX-IIA) (Hemolysin IIA) (HLY-IIA)
DE   (Cytolysin IIA) (CLY-IIA).
GN   APXIIA OR CLYIIA OR HLYIIA OR APPA OR CYTC.
OS   Actinobacillus pleuropneumoniae (Haemophilus pleuropneumoniae).
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC   Pasteurellaceae; Actinobacillus.
OX   NCBI_TaxID=715;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Serotype 5;
RX   MEDLINE=90126233; PubMed=2693022;
RA   Chang Y.-F., Young R., Struck D.K.;
RT   "Cloning and characterization of a hemolysin gene from Actinobacillus
RT   (Haemophilus) pleuropneumoniae.";
RL   DNA 8:635-647(1989).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Isolate CVI 13261 / Serotype 9;
RX   MEDLINE=92040145; PubMed=1937809;
RA   Smits M.A., Briaire J., Jansen R., Smith H.E., Kamp E.M.,
RA   Gielkens A.L.;
RT   "Cytolysins of Actinobacillus pleuropneumoniae serotype 9.";
RL   Infect. Immun. 59:4497-4504(1991).
CC   -!- FUNCTION: One of the virulence factors of A.pleuropneumoniae,
CC       which shows a weak hemolytic activity and is moderately cytotoxic
CC       for alveolar macrophages and neutrophils.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- DOMAIN: The Gly-rich region is probably involved in binding
CC       calcium, which is required for target cell-binding or cytolytic
CC       activity (By similarity).
CC   -!- DOMAIN: The three transmembrane domains are believed to be
CC       involved in pore formation by the cytotoxin (By similarity).
CC   -!- PTM: Palmitoylated by apxIIC. The toxin only becomes active when
CC       modified (By similarity).
CC   -!- SIMILARITY: BELONGS TO THE RTX PROKARYOTIC TOXIN FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; M30602; AAA87232.1; -.
DR   EMBL; X61111; CAA43423.1; -.
DR   PIR; B33389; B33389.
DR   InterPro; IPR001343; Hemlysn_Ca_bind.
DR   InterPro; IPR003995; RtxA.
DR   Pfam; PF00353; hemolysinCabind; 5.
DR   Pfam; PF02382; RTX; 1.
DR   PRINTS; PR00313; CABNDNGRPT.
DR   PRINTS; PR01488; RTXTOXINA.
DR   PROSITE; PS00330; HEMOLYSIN_CALCIUM; 1.
KW   Hemolysis; Toxin; Cytolysis; Repeat; Calcium; Transmembrane;
KW   Lipoprotein; Palmitate.
FT   TRANSMEM     233    256       POTENTIAL.
FT   TRANSMEM     266    323       POTENTIAL.
FT   TRANSMEM     361    406       POTENTIAL.
FT   DOMAIN       719    787       8 X REPEATS, GLY-RICH.
FT   REPEAT       719    724       1.
FT   REPEAT       728    733       2.
FT   REPEAT       737    742       3.
FT   REPEAT       746    751       4.
FT   REPEAT       755    760       5.
FT   REPEAT       764    769       6.
FT   REPEAT       773    778       7.
FT   REPEAT       782    787       8.
SQ   SEQUENCE   956 AA;  102531 MW;  BDBCABBADF14A641 CRC64;

  Query Match             48.1%;  Score 2236.5;  DB 1;  Length 956;
  Best Local Similarity   47.7%;  Pred. No. 4.4e-105;
  Matches  453;  Conservative  188;  Mismatches  255;  Indels   53;  Gaps   15;

Qy       1 MSNINV--IKSNIQAGL---------------------NSTKSGLKNLYLAIPKDYDPQ 36
Db       1 MSKITLSSLKSSLQQGLKNGKNKLNQAGTTLKNGLTQTGHSLQNGAKKLILYIPQGYDSG 60

Qy      37 KGGTLNDFIKAADELGIARLAEEPNHTETAKKSVDTVNQFLSLTQTGIAISATKLEKFLQ 96
Db      61 QGNGVQDLVKAANDLGIEVWREERSNLDIAKTSFDTTQKILGFTDRGIVLFAPQLDNLLK 120
```

```
Qy        97 KHSTNKLAKGLDSVENIDRKLGKASNVLSTLSSFLGTALAGIELDSLIKKGDAAPDALAK 156
             |:   |:   | |  :| : :|||: ||   : | ||: |:|: |: |::  |      |||
Db       121 KNP--KIGNTLGSASSISQNIGKANTVLGGIQSILGSVLSGVNLNELLQNKDPNQLELAK 178

Qy       157 ASIDLINEIIGNLSQSTQTIEAFSSQLAKLGSTISQAKGFSNIGNKLQNL-NFSKTNLGL 215
             | ::| ||::||:: | ||::||: |::|||  :    ||     : ||||||  :   | :|||
Db       179 AGLELTNELVGNIASSVQTVDAFAEQISKLGSHLQNVKGLGGLSNKLQNLPDLGKASLGL 238

Qy       216 EIITGLLSGISAGFALADKNASTGKKVAAGFELSNQVIGNVTKAISSYVLAQRVAAGLST 275
             :||:||||| |||  ||||  |||  ||  |||  |  :||:|||||||:|||:||||||:|||:
Db       239 DIISGLLSGASAGLILADKEASTEKKAAAGVEFANQIIGNVTKAVSSYILAQRVASGLSS 298

Qy       276 TGAVAALITSSIMLAISPLAFMNAADKFNHANALDEFAKQFRKFGYDGDHLLAEYQRGVG 335
             || |||||  |::  ||:|||:|:|  ||||  |: :   ::::|:|  |||||  |||::  |   |
Db       299 TGPVAALIASTVALAVSPLSFLNVADKFKQADLIKSYSERFQKLGYDGDRLLADFHRETG 358

Qy       336 TIEASLTTISTALGAVSAGVSAAAVGSAVGTPIALLVAGVTGLISGILEASKQAMFESVA 395
             ||:||:|||:||| |:|  ||  ||:  ||  ||  |:||||||||||||:  |||  ||||||  ||
Db       359 TIDASVTTINTALAAISGGVGAASAGSLVGAPVALLVAGVTGLITTILEYSKQAMFEHVA 418

Qy       396 NRLQGKILEWEKQNGGQNYFDKGYDSRYAAYLANNLKFLSELNKELEAERVIAITQQRWD 455
             |::   :|:||||::   :|||::|||||:  |  |  :|:|||   |||||:||||:|||||||
Db       419 NKVHDRIVEWEKKH-NKNYFEQGYDSRHLADLQDNMKFLINLNKELQAERVVAITQQRWD 477

Qy       456 NNIGELAGITKLGERIKSGKAYADAFEDGKKVEAGSNITLDAKTGIIDISNSNGKKTQAL 515
             | ||:||  |::    ::|  |||||  ||||:|            |::  ||  |  |||:|||:|  :|||::
Db       478 NQIGDLAAISRRTDKISSGKAYVDAFEEGQHQSYDSSVQLDNKNGIINISNTN-RKTQSV 536

Qy       516 HFTSPLLTAGTESRERLTNGKYSYINKLKFGRVKNWQVTDGEASSKLDFSKVIQRVA--- 572
              | :|||| | |:|||:  || ||| ||   ||  :| ||||:|||  :||:  |:||:|
Db       537 LFRTPLLTPGEENRERIQEGKNSYITKLHIQRVDSWTVTDGDASSSVDFTNVVQRIAVKF 596

Qy       573 -------ETEGTDEIGLIVNAKAGNDDIFVGQGKMNIDGGDGHDRVFYSKDGGFGNITVD 625
                    |::  |      :|  |   ||||::|||    |||||||||  ||:  |  :|  :  :|
Db       597 DDAGNIIESKDTK---IIANLGAGNDNVFVGSSTTVIDGGDGHDRVHYSR-GEYGALVID 652

Qy       626 GTSATEAGSYTVNRKVARGDIYHEVVKRQETKVGKRTETIQYRDYELRKVGYGYQSTDNL 685
              |: || |||:| |  |         || :     :| ||  |  |  | |:||    |    :    ||    ||:|
Db       653 ATAETEKGSYSVKRYVGDSKALHETIATHQTNVGNREEKIEYR-REDDRFHTGYTVTDSL 711

Qy       686 KSVEEVIGSQFNDVFKGSKFNDIFHSGEGDDLLDGGAGDDRLFGGKGNDRLSGDEGDDLL 745
             |||||:||||||||:|||||:|:|:||  |  |  |  |  :|   |||   ||||   |:|    :   |       |::  |
Db       712 KSVEEIIGSQFNDIFKGSQFDDVFHGGNGVDTIDGNDGDDHLFGGAGDDVIDGGNGNNFL 771

Qy       746 DGGSGDDVLNGGAGNDVYIFRKGDGNDTLYDGTGNDKLAFADANISDIMIERTKEGIIVK 805
             ||:|:|:::||   ||:|:  :  ||||::   |    ||||:|  |: |:  ::      : :
Db       772 VGGTGNDIISGGKDNDIYVHKTGDGNDSITDSGGQDKLAFSDVNLKDLTFKKVDSSLEI- 830

Qy       806 RNDHSGSINIPRWY----ITSNLQNYQSNKTDHKIEQLIGKDGSYITSDQIDKILQDKKD 861
              |      :  |   |:      :  |  :  ||::     |  |||::|||  |    |||:|:||::   |:
Db       831 INQKGEKVRIGNWFLEDDLASTVANYKAT-NDRKIEEIIGKGGERITSEQVDKLI--KEG 887

Qy       862 GTVITSQELKKLADENKSQKLSASDIASSLNKLVGSMALFGTANSVSSN 910
                 |:::  |  |: ::   :  |       ::::||  ||:  |:    |  :::    :|
Db       888 NNQISAEALSKVVNDYNTSK-DRQNVSNSLAKLISSVGSFTSSSDFRNN 935
```

RESULT 6
LKAA_PASHA
ID   LKAA_PASHA   STANDARD;   PRT;   955 AA.
AC   P55117;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Leukotoxin from serotype T10.
GN   LKTA.
OS   Pasteurella haemolytica.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC   Pasteurellaceae; Mannheimia.
OX   NCBI_TaxID=75985;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Serotype T10;
RX   MEDLINE=96425875; PubMed=8828217;
RA   Lainson F.A., Murray J., Davies R.C., Donachie W.;
RT   "Characterization of epitopes involved in the neutralization of
RT   Pasteurella haemolytica serotype A1 leukotoxin.";
RL   Microbiology 142:2499-2507(1996).
CC   -!- FUNCTION: Bacterial hemolysins are exotoxins that attack blood
CC   cell membranes and cause cell rupture by mechanisms not clearly
CC   defined.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- DOMAIN: The Gly-rich region is probably involved in binding
CC   calcium, which is required for target cell-binding or cytolytic
CC   activity.
CC   -!- DOMAIN: The three transmembrane domains are believed to be
CC   involved in pore formation by the cytotoxin (BY SIMILARITY).
CC   -!- PTM: Palmitoylated by lktC. The toxin only becomes active when
CC   MODIFIED (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE RTX PROKARYOTIC TOXIN FAMILY.
CC   --

CC   --
DR   EMBL; Z26247; CAA81206.1; -.
DR   PIR; S37145; A35254.
DR   InterPro; IPR001343; Hemlysn_Ca_bind.
DR   InterPro; IPR003995; RtxA.
DR   Pfam; PF00353; hemolysinCabind; 5.
DR   Pfam; PF02382; RTX; 1.
DR   PRINTS; PR00313; CABNDNGRPT.
DR   PRINTS; PR01488; RTXTOXINA.
DR   PROSITE; PS00330; HEMOLYSIN_CALCIUM; 4.
KW   Hemolysis; Toxin; Cytolysis; Repeat; Calcium; Transmembrane;
KW   Lipoprotein; Palmitate.

FT	TRANSMEM	299	319	POTENTIAL.
FT	TRANSMEM	361	381	POTENTIAL.
FT	TRANSMEM	383	403	POTENTIAL.
FT	DOMAIN	736	786	6 X REPEATS, GLY-RICH.
FT	REPEAT	736	741	1.
FT	REPEAT	745	750	2.
FT	REPEAT	754	759	3.
FT	REPEAT	763	768	4.
FT	REPEAT	772	777	5.
FT	REPEAT	781	786	6.

SQ   SEQUENCE   955 AA; 102187 MW; B60F2DB8168EBCAF CRC64;

Query Match 48.1%; Score 2236; DB 1; Length 955;
Best Local Similarity 48.2%; Pred. No. 4.7e-105;
Matches 458; Conservative 178; Mismatches 271; Indels 44; Gaps 17;

```
Qy         5 NVIKSNIQA---GLN-----------STKSGLKNLYLAIPKD--YDPQKGGTLNDFIKAA 48
             | |:|  :  |    |||               :  |:|  |  :  |  ||||  ||    |   |  :|||
Db        12 NGIRSTLTATRGGLNRAGQSLTQAGQTLKNGAKKIILYIPKDYKYDSGSGNGLQDLVKAA 71

Qy        49 DELGIARLAEEPNHTETAKKSVDTVNQFLSLTQTGIAISATKLEKFLQKHSTNKLAKGLD 108
             :||||     || |      |: |: |:     | ||: || :|| :|:| |||      ||: :  |
Db        72 EELGIEVQKEEGNDIAKAQTSLGTIQNVLGLTERGIVLSAPQLDKLLQK---NKVGQALG 128

Qy       109 SVENIDRKLGKASNVLSTLSSFLGTALAGIELDSLIKKGDAAPDALAKASIDLINEIIGN 168
             | |:  :   :|  |||  :        | |||::||  : : ::          |||| ::|  :|  |
Db       129 SSESTAQNFSQAKTVLSGVQGNSRTVLAGMDLDEAL-QNESDQLTLAKAGLELTNSLIEN 187

Qy       169 LSQSTQTIEAFSSQLAKLGSTISQAKGFSNIGNKLQNL-NFSKTNLGLEIITGLLSGISA 227
             :: | ||::||| |::: ||  :    ||       :|:||:|:      |   |||:: : |||| ::
Db       188 IANSVQTLDAFSEQISQFGSKLQNVKGLGALGDKLKNIGGLDKAGLGLDVKSRLLSGATA 247

Qy       228 GFALADKNASTGKKVAAGFELSNQVIGNVTKAISSYVLAQRVAAGLSTTGAVAALITSSI 287
                  ||||:||| |||  |||||:|||:|||:||||:||||||||||||:|| ||||| ::
Db       248 ALVLADKDASTAKKVGAGFELANQVVGNITKAVSSYILAQRVAAGLSSTGPVAALIASTV 307
```

```
Qy        288 MLAISPLAFMNAADKFNHANALDEFAKQFRKFGYDGDHLLAEYQRGVGTIEASLTTISTA 347
               :|||||:|    ||||: | :|: :|::|:| ||:|| |||||| | |||:||:| |:||
Db        308 AVAISPLSFAGIADKFDRAKSLENYAERFKKLGYEGDSLLAEYQHGTGTIDASVTAINTA 367

Qy        348 LGAVSAGVSAAAVGSAVGTPIALLVAGVTGLISGILEASKQAMFESVANRLQGKILEWEK 407
              | |:: |||||| || | :||||||:|:|:|| ||: |||||| |||::  ||:||||
Db        368 LAAIAGGVSAAAAGSVVASPIALLVSGITGVISTILQYSKQAMFEHVANKIHNKIVEWEK 427

Qy        408 QNGGQNYFDKGYDSRYAAYLANNLKFLSELNKELEAERVIAITQQRWDNNIGELAGITKL 467
               |||:|||: |||:|| | | :|:||| |||||:||||||||||:|:|||:||||::|
Db        428 NNGGKNYFENGYDARYLANLQDNMKFLLNLNKELQAERVIAITQQQWDSNIGDLAGISRL 487

Qy        468 GERIKSGKAYADAFEDGKKVEAGSNITLDAKTGIIDISNSNGKKTQALHFTSPLLTAGTE 527
              ||:: ||||| ||||:|: ::|     : ||:  ||||::|:    ||| : | :|||| |||
Db        488 GEKVLSGKAYVDAFEEGQHLKADKLVQLDSAKGIIDVTNTGEAKTQHILFRTPLLTPGTE 547

Qy        528 SRERLTNGKYSYINKLKFGRVKNWQVTDGEASSKLDFSKVIQRVA-------ETEGTDEI 580
               |||:  ||| || ||    || :||: || |||   | : |:||:           | |
Db        548 KRERVQTGKYEYITKLHINRVDSWQIKDGAASSTFDLTNVVQRIGVELDHAENVIKTKET 607

Qy        581 GLIVNAKAGNDDIFVGQGKMNIDGGDGHDRVFYSKDGGFGNITVDGTSATEAGSYTVNRK 640
               ::       |:|::|||  |    ||||:|:|||  ||: | :| :|:| |   ||||||
Db        608 KIVATLGDGDDNVFVGSGTTEIDGGEGYDRVHYSR-GNYGALTIDATKETEQGSYTVNRF 666

Qy        641 VARGDIYHEVVKRQETKVGKRTETIQYRDYELRKVGYGYQSTDNLKSVEEVIGSQFNDVF 700
              |   |      ||                   ||  |  :   || : | ||:||||:  ||:|
Db        667 VESGKALHEGTSTHTALVGNREEKIEYR-HSNNQHHAGYYTKDTLKAVEEIIGTSHNDIF 725

Qy        701 KGSKFNDIFHSGEGDDLLDGGAGDDRLFGGKGNDRLSGDEGDDLLDGGSGDDVLNGGAGN 760
              |||||| |: |:| | :||  |:|||||||:| : | ||| :||| |:|:|:|| |:
Db        726 KGSKFNDAFNGGDGVDTIDGNDGNDRLFGGKGDDIIDGGNGDDFIDGGKGNDLLHGGKGD 785

Qy        761 DVYIFRKGDGNDTLYDGTGNDKLAFADANISDIMIERTKEGIIVKRNDHSGSINIPRWY- 819
              |::: |:||||||:: :   |||||:|:|:|:|: |: |:      :::  |        : | |:
Db        786 DIFVHRQGDGNDSITESEGNDKLSFSDSNLKDLTFEKVNHHLVI-TNTKQEKVTIQNWFR 844

Qy        820 ---ITSNLQNYQSNKTDHKIEQLIGKDGSYITSDQIDKILQDKKDGTVITSQELKKLADE 876
                  :||| : :  | |||::||::|  |||  |:|::::  |     |    || |: |
Db        845 EAEFAKTIQNYVATR-DDKIEEIIGQNGERITSKQVDELIE--KGNGKIAQSELTKVVDN 901

Qy        877 NKSQKLSASDIASSLNKLVGSMALFGTANSVSSNALQPITQPTQGILAPSV 927
               :  | |  | ::||:||: | :  | ::|  | | |      ||  :| ||:
Db        902 YQLLKYS-RDASNSLDKLISSASAFTSSND-SRNVL---ASPT-SMLDPSL 946
```

RESULT 7
HLYA_ACTSU

```
ID   HLYA_ACTSU     STANDARD;      PRT;   956 AA.
AC   Q00951;
DT   01-APR-1993 (Rel. 25, Created)
DT   01-APR-1993 (Rel. 25, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Hemolysin (Cytolysin II) (CLY-IIA) (HLY-IIA) (CYTC) (APPA).
GN   APPA OR CLYIIA OR HLYIIA OR CYTC.
OS   Actinobacillus suis.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC   Pasteurellaceae; Actinobacillus.
OX   NCBI_TaxID=716;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=3714;
RX   MEDLINE=92267623; PubMed=1587585;
RA   Burrows L.L., Lo R.Y.;
RT   "Molecular characterization of an RTX toxin determinant from
RT   Actinobacillus suis.";
RL   Infect. Immun. 60:2166-2173(1992).
CC   -!- FUNCTION: One of the virulence factors of A.suis might be a
CC       secreted cytotoxin, possibly the extracellular hemolysin.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- DOMAIN: The Gly-rich region is probably involved in binding
CC       calcium, which is required for target cell-binding or cytolytic
CC       activity.
CC   -!- DOMAIN: The three transmembrane domains are believed to be
CC       involved in pore formation by the cytotoxin (By similarity).
CC   -!- PTM: Palmitoylated by lktC. The toxin only becomes active when
CC       modified (By similarity).
CC   -!- SIMILARITY: BELONGS TO THE RTX PROKARYOTIC TOXIN FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; M90440; AAA21918.1; -.
DR   InterPro; IPR001343; Hemlysn_Ca_bind.
DR   InterPro; IPR003995; RtxA.
DR   Pfam; PF00353; hemolysinCabind; 5.
DR   Pfam; PF02382; RTX; 1.
DR   PRINTS; PR00313; CABNDNGRPT.
DR   PRINTS; PR01488; RTXTOXINA.
DR   PROSITE; PS00330; HEMOLYSIN_CALCIUM; 1.
KW   Hemolysis; Toxin; Cytolysis; Repeat; Calcium; Transmembrane;
KW   Lipoprotein; Palmitate.
FT   TRANSMEM    238    254       POTENTIAL.
FT   TRANSMEM    302    320       POTENTIAL.
FT   TRANSMEM    383    406       POTENTIAL.
FT   DOMAIN      719    799       9 X REPEATS, GLY-RICH.
FT   REPEAT      719    724       1.
FT   REPEAT      728    733       2.
FT   REPEAT      737    742       3.
FT   REPEAT      746    751       4.
FT   REPEAT      755    760       5.
FT   REPEAT      764    769       6.
FT   REPEAT      773    778       7.
FT   REPEAT      782    787       8.
FT   REPEAT      794    799       9.
SQ   SEQUENCE   956 AA;  102453 MW;  3415FF1D7ADD4365 CRC64;

  Query Match             47.7%;  Score 2214.5;  DB 1;  Length 956;
  Best Local Similarity   47.4%;  Pred. No. 5.6e-104;
  Matches  450;  Conservative  188;  Mismatches  258;  Indels   53;  Gaps   15;

Qy          1 MSNINV--IKSNIQAGL---------------------NSTKSGLKNLYLAIPKDYDPQ 36
              || |  :  :||::| ||                     :| ::| | | | ||: ||
Db          1 MSKITLSSLKSSLQQGLKNGKNKLNQAGTTLKNGLTQTGHSLQNGAKKLILYIPQGYDSG 60

Qy         37 KGGTLNDFIKAADELGIARLAEEPNHTETAKKSVDTVNQFLSLTQTGIAISATKLEKFLQ 96
               :|   :  | :|||::|||     ||    |     |    ||  :   | | :|: |:
Db         61 QGNGIQDLVKAANDLGIEVWREERSNLDIAKTSFDTTQKILGFTDRGIVLFAPQLDNLLK 120

Qy         97 KHSTNKLAKGLDSVENIDRKLGKASNVLSTLSSFLGTALAGIELDSLIKKGDAAPDALAK 156
              |:   |:    |  |  :| : :|||: ||   :  | ||: |:|: |: |::   |    |||
Db        121 KNP--KIGNTLGSASSISQNIGKANTVLGGIQSILGSVLSGVNLNELLQNKDPNQLELAK 178

Qy        157 ASIDLINEIIGNLSQSTQTIEAFSSQLAKLGSTISQAKGFSNIGNKLQNL-NFSKTNLGL 215
               | ::| ||::||:: | ||::||: |::|||  :     ||    : |||||  :   | :|||
Db        179 AGLELTNELVGNIASSVQTVDAFAEQISKLGSHLQNVKGLGGLSNKLQNLPDLGKASLGL 238

Qy        216 EIITGLLSGISAGFALADKNASTGKKVAAGFELSNQVIGNVTKAISSYVLAQRVAAGLST 275
               :||:||||| |||  |||| ||| || ||| | :||:||||||:|||:||||||:|||:
Db        239 DIISGLLSGASAGLILADKKASTEKKAAAGVEFANQIIGNVTKAVSSYILAQRVASGLSS 298

Qy        276 TGAVAALITSSIMLAISPLAFMNAADKFNHANALDEFAKQFRKFGYDGDHLLAEYQRGVG 335
              || |||| |::  ||:|||:|:| ||||   |:  :  ::::|:| |||| |||:: |
Db        299 TGPVAALIASTVALAVSPLSFLNVADKFKQADLIKSYSERFQKLGYDGDRLLADFHRETG 358

Qy        336 TIEASLTTISTALGAVSAGVSAAAVGSAVGTPIALLVAGVTGLISGILEASKQAMFESVA 395
              ||:||:|||:||| |:| || ||: || || |:||||||||||: ||| ||||||| ||
Db        359 TIDASVTTINTALAAISGGVGAASAGSLVGAPVALLVAGVTGLITTILEYSKQAMFEHVA 418

Qy        396 NRLQGKILEWEKQNGGQNYFDKGYDSRYAAYLANNLKFLSELNKELEAERVIAITQQRWD 455
              |::  :|:||||::   :|||::||||: | | :|:||| ||||||:||||:||||||||
Db        419 NKVHDRIVEWEKKH-NKNYFEQGYDSRHLADLQDNMKFLINLNKELQAERVVAITQQRWD 477
```

```
Qy        456 NNIGELAGITKLGERIKSGKAYADAFEDGKKVEAGSNITLDAKTGIIDISNSNGKKTQAL 515
              | ||:|| |::  ::| ||||| ||||:|      :      ::| | |||:|||:| :|||::
Db        478 NQIGDLAAISRRTDKISSGKAYVDAFEEGNTSPSIHPYSIDNKNGIINISNTN-RKTQSV 536

Qy        516 HFTSPLLTAGTESRERLTNGKYSYINKLKFGRVKNWQVTDGEASSKLDFSKVIQRVA--- 572
                |:|||| | |:|||: || ||| ||     || :| || |:||| :||: |:||:|
Db        537 LFRTPLLTPGEENRERIQEGKNSYITKLHIQRVDSWTVTVGDASSSVDFTNVVQRIAVKF 596

Qy        573 -------ETEGTDEIGLIVNAKAGNDDIFVGQGKMNIDGGDGHDRVFYSKDGGFGNITVD 625
                     |:: |    :| |  ||||::|||     ||||||||| ||: | :| : :|
Db        597 DDAGNIIESKDTK---IIANLGAGNDNVFVGSSTTVIDGGDGHDRVHYSR-GEYGALVID 652

Qy        626 GTSATEAGSYTVNRKVARGDIYHEVVKRQETKVGKRTETIQYRDYELRKVGYGYQSTDNL 685
               |: || |||:| |         || :   :| || | | |:||   |  :    || ||:|
Db        653 ATAETEKGSYSVKRYVGDSKALHETIATHQTNVGNREEKIEYR-REDDRFHTGYTVTDSL 711

Qy        686 KSVEEVIGSQFNDVFKGSKFNDIFHSGEGDDLLDGGAGDDRLFGGKGNDRLSGDEGDDLL 745
              |||||:|||||||:||||:|:|:|| | | | :||   ||| |||| |:| : |   |:: |
Db        712 KSVEEIIGSQFNDIFKGSQFDDVFHGGNGVDTIDGNDGDDHLFGGAGDDVIDGGNGNNFL 771

Qy        746 DGGSGDDVLNGGAGNDVYIFRKGDGNDTLYDGTGNDKLAFADANISDIMIERTKEGIIVK 805
               ||:|:|:::|| ||:|:  ||:|:  | |||||:| |: |:  ::     :
Db        772 VGGTGNDIISGGKDNDIYVHKTGDGNDSITDSGGQDKLAFSDVNLKDLTFKKVDSSLEI- 830

Qy        806 RNDHSGSINIPRWY----ITSNLQNYQSNKTDHKIEQLIGKDGSYITSDQIDKILQDKKD 861
                |    : |  |:     : | : ||::   | |||::||| |   |||:|:||:: |:
Db        831 INQKGEKVRIGNWFLEDDLASTVANYKAT-NDRKIEEIIGKGGERITSEQVDKLI--KEG 887

Qy        862 GTVITSQELKKLADENKSQKLSASDIASSLNKLVGSMALFGTANSVSSN 910
                  |::: | |: ::  : |     ::::|| ||: |:  | :::   :|
Db        888 NNQISAEALSKVVNDYNTSK-DRQNVSNSLAKLISSVGSFTSSSDFRNN 935
```

```
RESULT 8
RT32_ACTPL
ID   RT32_ACTPL     STANDARD;      PRT;  1052 AA.
AC   P55131;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   RTX-III toxin determinant A from serotype 8 (APX-IIIA) (Cytolysin
DE   IIIA) (CLY-IIIA).
GN   APXIIIA OR CLYIIIA OR RTXA OR PTXA.
OS   Actinobacillus pleuropneumoniae (Haemophilus pleuropneumoniae).
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC   Pasteurellaceae; Actinobacillus.
OX   NCBI_TaxID=715;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=405 / Serotype 8;
RX   MEDLINE=95012630; PubMed=7927703;
RA   Jansen R., Briaire J., van Geel A.B.M., Kamp E.M., Gielkens A.L.J.,
RA   Smits M.A.;
RT   "Genetic map of the Actinobacillus pleuropneumoniae RTX-toxin (Apx)
RT   operons: characterization of the ApxIII operons.";
RL   Infect. Immun. 62:4411-4418(1994).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Serotype 8;
RX   MEDLINE=93162836; PubMed=8432615;
RA   Jansen R., Briaire J., Kamp E.M., Gielkens A.L.J., Smits M.A.;
RT   "Cloning and characterization of the Actinobacillus
RT   pleuropneumoniae-RTX-toxin III (ApxIII) gene.";
RL   Infect. Immun. 61:947-954(1993).
CC   -!- FUNCTION: Does not have hemolytic activity but shows a strong
CC       cytotoxicity towards alveolar macrophages and neutrophils.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- DOMAIN: The Gly-rich region is probably involved in binding
CC       calcium, which is required for target cell-binding or cytolytic
CC       activity (By similarity).
CC   -!- DOMAIN: The three transmembrane domains are believed to be
CC       involved in pore formation by the cytotoxin (By similarity).
CC   -!- PTM: Palmitoylated by apxIIIC. The toxin only becomes active when
CC       modified (By similarity).
CC   -!- SIMILARITY: BELONGS TO THE RTX PROKARYOTIC TOXIN FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; X80055; CAA56358.1; -.
DR   EMBL; X68815; CAA48711.1; -.
DR   PIR; B49219; B49219.
DR   InterPro; IPR001343; Hemlysn_Ca_bind.
DR   InterPro; IPR003995; RtxA.
DR   Pfam; PF00353; hemolysinCabind; 6.
DR   Pfam; PF02382; RTX; 1.
DR   PRINTS; PR00313; CABNDNGRPT.
DR   PRINTS; PR01488; RTXTOXINA.
DR   PROSITE; PS00330; HEMOLYSIN_CALCIUM; 3.
KW   Toxin; Cytolysis; Repeat; Calcium; Transmembrane; Lipoprotein;
KW   Palmitate.
FT   TRANSMEM     248    265       POTENTIAL.
FT   TRANSMEM     275    334       POTENTIAL.
FT   TRANSMEM     372    418       POTENTIAL.
FT   DOMAIN       754    859       7 X REPEATS, GLY-RICH.
FT   REPEAT       754    759       1.
FT   REPEAT       763    768       2.
FT   REPEAT       772    777       3.
FT   REPEAT       781    786       4.
FT   REPEAT       790    795       5.
FT   REPEAT       799    804       6.
FT   REPEAT       808    813       7.
FT   REPEAT       827    832       8.
FT   REPEAT       836    841       9.
FT   REPEAT       845    850       10.
FT   REPEAT       854    859       11.
SQ   SEQUENCE   1052 AA;  112809 MW;  F83AFE25A6FD8758 CRC64;

  Query Match             42.1%;  Score 1955.5;  DB 1;  Length 1052;
  Best Local Similarity   42.6%;  Pred. No. 6.6e-91;
  Matches  425;  Conservative  166;  Mismatches  287;  Indels  119;  Gaps   19;

Qy          8 KSNIQAGLNSTKSGLK---------------------NLYLAIPKDYDPQKGGTLNDFI 45
              |   : | : ||:||:                      | | |||:||    |     | :
Db         19 KRQAKKGYDVTKNGLQYGVSQAKLQALAAGKAVQKYGNKLVLVIPKEYDGSVGNGFFDLV 78

Qy         46 KAADELGIARLAEEPNHTETAKKSVDTVNQFLSLTQTGIAISATKLEKFLQKHSTNKLAK 105
              |||:||||        |  | | ||: | :||| ||: |: :  | :|::||||||
Db         79 KAAEELGIQVKYVNRNELEVAHKSLGTADQFLGLTERGLTLFAPQLDQFLQKHSKISNVV 138

Qy        106 GLDSVENIDRKLGKASNVLSTLSSFLGTALAGIELDSLIKKGDAAPDALAKASIDLINEI 165
              |  : :  :   || |:   ::| : | ||| ||||| |:  |   |  :   ||:| : | :|:
Db        139 GSSTGDAVS-KLAKSQTIISGIQSVLGTVLAGINLNEAIISGGSELE-LAEAGVSLASEL 196

Qy        166 IGNLSQSTQTIEAFSSQLAKLGSTISQAKGFSNIGNKLQNLN---FSKTNLGLEIITGLL 222
              : |::: | ||:|::|:     |   :  |||     :| :|||::      ||| |||:||: ||
Db        197 VSNIAKGTTTIDAFTTQIQNFGKLVENAKGLGGVGRQLQNISGGALSKTGLGLDIISSLL 256

Qy        223 SGISAGFALADKNASTGKKVAAGFELSNQVIGNVTKAISSYVLAQRVAAGLSTTGAVAAL 282
              ||::| ||||:||||   |||||||||||| :|||:|||:||||:||||||||   |||
Db        257 SGVTASFALANKNASTSTKVAAGFELSNQVIGGITKAVSSYILAQRLAAGLSTTGPAAAL 316

Qy        283 ITSSIMLAISPLAFMNAADKFNHANALDEFAKQFRKFGYDGDHLLAEYQRGVGTIEASLT 342
              | ||| ||||||||:  || ||  :   : |||::|:| ||||| ||:|:       |||:||:|
Db        317 IASSISLAISPLAFLRVADNFNRSKEIGEFAERFKKLGYDGDKLLSEFYHEAGTIDASIT 376

Qy        343 TISTALGAVSAGVSAAAVGSAVGTPIALLVAGVTGLISGILEASKQAMFESVANRLQGKI 402
              |||||| |::|| :||: |: || |||| |:||||||||| ||| |  : ||:::  ||
Db        377 TISTALSAIAAGTAAASAGALVGAPITLLVTGITGLISGILEFSKQPMLDHVASKIGNKI 436
```

```
Qy        403 LEWEKQNGGQNYFDKGYDSRYAAYLANNLKFLSELNKELEAERVIAITQQRWDNNIGELA 462
              ||||:  |:|||: |||:|: |:| ::  ||  ||: | || : ||||||| |||||
Db        437 DEWEKKY-GKNYFENGYDARHKAFLEDSFSLLSSFNKQYETERAVLITQQRWDEYIGELA 495

Qy        463 GITKLGERIKSGKAYADAFEDGKKVEAG----SNITLDAKTGIIDISNSNGKKTQALHFT 518
              |||  |::: ||||| | |::|| :|         | :   |  ||||||   : :  | |
Db        496 GITGKGDKLSSGKAYVDYFQEGKLLEKKPDDFSKVVFDPTKGEIDISNS--QTSTLLKFV 553

Qy        519 SPLLTAGTESRERLTNGKYSYINKLKFGRVKNWQVTDG--EASSKLDFSKVIQR------ 570
              :|||| |||||||  || ||        | | :|  :  :   |:: :||
Db        554 TPLLTPGTESRERTQTGKYEYITKLVVKGKDKW-VVNGVKDKGAVYDYTNLIQHAHISSS 612

Qy        571 VAETEGTDEIGLIVNAKAGNDDIFVGQGKMNIDGGDGHDRVFYSKDGGFGNITVDGTSAT 630
              || |    |: |: :   ||| :|:  |   |   |:||| |:| |   | : :||| ||
Db        613 VARGEEYREVRLVSHLGNGNDKVFLAAGSAEIHAGEGHDVVYYDKT-DTGLLVIDGTKAT 671

Qy        631 EAGSYTVNRKVARG-DIYHEVVKRQETKVGKRTETIQYRDYELRKVG-YGYQSTDNLKSV 688
              | | |:| |:::     |  ||:| |:: |||| ||::|||||| : |     :: | | ||
Db        672 EQGRYSVTRELSGATKILREVIKNQKSAVGKREETLEYRDYELTQSGNSNLKAHDELHSV 731

Qy        689 EEVIGSQFNDVFKGSKFNDIFHSGEGDDLLDGGAGDDRLFGGKGNDRLSGDEGD------ 742
              ||:|||   |  ||||||| ||||    :|||||:|  |||  ||:|  ||||  || ||  |:
Db        732 EEIIGSNQRDEFKGSKFRDIFHGADGDDLLNGNDGDDILYGDKGNDELRGDNGNDQLYGG 791

Qy        743 ------------------------------------------------DLLDGGSGDDV 753
                                                              |||||| |:|
Db        792 EGNDKLLGGNGNNYLSGGDGNDELQVLGNGFNVLRGGKGDDKLYGSSGSDLLDGGEGNDY 851

Qy        754 LNGGAGNDVYIFRKGDGNDTLYD---GTGNDKLAFADANISDIMIERTKEGIIVKRNDHS 810
              | || |:| |::|   || |:||   :  |||  :| :    :::|:  : :::: |: |
Db        852 LEGGDGSDFYVYRSTSGNHTIYDQGKSSDLDKLYLSDFSFDRLLVEKVDDNLVLRSNESS 911

Qy        811 ---GSINIPRWYITSNLQNYQSNKTDHKIEQLIGKDGSYITSDQIDKILQDKKDGTVITS 867
                 | : |  |:          : || :|||||:: |:|  :|:: :     ::      : :
Db        912 HNNGVLTIKDWF-------KEGNKYNHKIEQIVDKNGRKLTAENLGTYFKNAPKADNLLN 964

Qy        868 QELKKLADENKSQKLSASDIASSLNKLVGSMALFGTA 904
                 |:  |:|:|   : | : : |:|:: :    || |
Db        965 YATKE--DQNES---NLSSLKTELSKIITNAGNFGVA 996
```

RESULT 9

RT31_ACTPL

```
ID   RT31_ACTPL     STANDARD;      PRT;  1049 AA.
AC   P55130;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   RTX-III toxin determinant A from serotype 2 (APX-IIIA) (Cytolysin
DE   IIIA) (CLY-IIIA).
GN   APXIIIA OR CLYIIIA OR RTXA OR PTXA.
OS   Actinobacillus pleuropneumoniae (Haemophilus pleuropneumoniae).
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC   Pasteurellaceae; Actinobacillus.
OX   NCBI_TaxID=715;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Serotype 2;
RX   MEDLINE=93263992; PubMed=8494611;
RA   Chang Y.-F., Shi J., Ma D.-P., Shin S.J., Lein D.H.;
RT   "Molecular analysis of the Actinobacillus pleuropneumoniae RTX
RT   toxin-III gene cluster.";
RL   DNA Cell Biol. 12:351-362(1993).
RN   [2]
RP   SEQUENCE OF 828-1049 FROM N.A.
RC   STRAIN=1536 / Serotype 2;
RX   MEDLINE=95012630; PubMed=7927703;
RA   Jansen R., Briaire J., van Geel A.B.M., Kamp E.M., Gielkens A.L.J.,
RA   Smits M.A.;
RT   "Genetic map of the Actinobacillus pleuropneumoniae RTX-toxin (Apx)
RT   operons: characterization of the ApxIII operons.";
RL   Infect. Immun. 62:4411-4418(1994).
CC   -!- FUNCTION: Does not have hemolytic activity but shows a strong
CC       cytotoxicity towards alveolar macrophages and neutrophils.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- DOMAIN: The Gly-rich region is probably involved in binding
CC       calcium, which is required for target cell-binding or cytolytic
CC       activity (By similarity).
CC   -!- DOMAIN: The three transmembrane domains are believed to be
CC       involved in pore formation by the cytotoxin (By similarity).
CC   -!- PTM: Palmitoylated by apxIIIC. The toxin only becomes active when
CC       modified (By similarity).
CC   -!- SIMILARITY: BELONGS TO THE RTX PROKARYOTIC TOXIN FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; L12145; AAA21924.1; -.
DR   EMBL; X80056; CAB37652.1; ALT_SEQ.
DR   PIR; S51784; S51784.
DR   InterPro; IPR001343; Hemlysn_Ca_bind.
DR   InterPro; IPR003995; RtxA.
DR   Pfam; PF00353; hemolysinCabind; 6.
DR   Pfam; PF02382; RTX; 1.
DR   PRINTS; PR00313; CABNDNGRPT.
DR   PRINTS; PR01488; RTXTOXINA.
DR   PROSITE; PS00330; HEMOLYSIN_CALCIUM; 3.
KW   Toxin; Cytolysis; Repeat; Calcium; Transmembrane; Lipoprotein;
KW   Palmitate.
FT   TRANSMEM    154    170       POTENTIAL.
FT   TRANSMEM    315    331       POTENTIAL.
FT   TRANSMEM    397    413       POTENTIAL.
FT   DOMAIN      753    858       11 X REPEATS, GLY-RICH.
FT   REPEAT      753    758       1.
FT   REPEAT      762    767       2.
FT   REPEAT      771    776       3.
FT   REPEAT      780    785       4.
FT   REPEAT      789    794       5.
FT   REPEAT      798    803       6.
FT   REPEAT      807    812       7.
FT   REPEAT      826    831       8.
FT   REPEAT      835    840       9.
FT   REPEAT      844    849       10.
FT   REPEAT      853    858       11.
SQ   SEQUENCE   1049 AA;  112491 MW;  F99846BFD4E5CE72 CRC64;

  Query Match             41.4%;  Score 1924;  DB 1;  Length 1049;
  Best Local Similarity   43.0%;  Pred. No. 2.5e-89;
  Matches  428;  Conservative  152;  Mismatches  296;  Indels  120;  Gaps   21;

Qy          8 KSNIQAGLNSTKSGLK---------------------NLYLAIPKDYDPQKGGTLNDFI 45
              |  ::  |  :  ||:||:                     | |  |||:||   |     |  :
Db         19 KRQVKKGYDVTKNGLQYGVSQAKLQALAAGKAVQKYGNKLVLVIPKEYDGSVGNGFFDLV 78

Qy         46 KAADELGIARLAEEPNHTETAKKSVDTVNQFLSLTQTGIAISATKLEKFLQKHSTNKLAK 105
              |||:|:|||         | |  | ||:   | :|||  ||:  |:  :  | :|:||||||
Db         79 KAAEELGIQVKYVNRNELEVAHKSLGTADQFLGLTERGLTLFAPQLDQFLQKHSKISNVV 138

Qy        106 GLDSVENIDRKLGKASNVLSTLSSFLGTALAGIELDSLIKKGDAAPDALAKASIDLINEI 165
              |   :  :  :  || |:  ::|  :  | ||| ||||  |:   |    |  :  :  ||:|  :  | :|:
Db        139 GSSTGDAVS-KLAKSQTIISGIQSVLGTVLAGINLNEAIISGGSELE-LAEAGVSLASEL 196

Qy        166 IGNLSQSTQTIEAFSSQLAKLGSTISQAKGFSNIGNKLQNLN---FSKTNLGLEIITGLL 222
              : |::: | ||:||::|:   |        |||     :| :|||::     |||  |||:||: ||
Db        197 VSNIAKGTTTIDAFTTQIQNFGKLAENAKGLGGVGRQLQNISGSALSKTGLGLDIISSLL 256

Qy        223 SGISAGFALADKNASTGKKVAAGFELSNQVIGNVTKAISSYVLAQRVAAGLSTTGAVAAL 282
              ||::  |||  :|||||  |||||||||||||  :|||:|||:||||:  |||||||  |||
Db        257 SGVTRSFALRNKNASTSTKVAAGFELSNQVIGGITKAVSSYILAQRLRAGLSTTGPAAAL 316

Qy        283 ITSSIMLAISPLAFMNAADKFNHANALDEFAKQFRKFGYDGDHLLAEYQRGVGTIEASLT 342
```

```
               | |||  ||||||||:  || ||  :   : |||::|:|  |||||  ||:|:      |||:||:|
Db        317 IASSISLAISPLAFLRVADNFNRSKEIGEFAERFKKLGYDGDKLLSEFYHEAGTIDASIT 376

Qy        343 TISTALGAVSAGVSAAAVGSAVGTPIALLVAGVTGLISGILEASKQAMFESVANRLQGKI 402
              ||||||  |::||  :||:  |:  ||  ||  |||  |:||||||||||  |||  |   :  ||:::   ||
Db        377 TISTALSAIAAGTAAASAGALVGAPITLLVTGITGLISGILEFSKQPMLDHVASKIGNKI 436

Qy        403 LEWEKQNGGQNYFDKGYDSRYAAYLANNLKFLSELNKELEAERVIAITQQRWDNNIGELA 462
               ||||:   |:|||:  |||:|:  |:|  ::      ||    ||:   | ||   :  |||||||   |||||
Db        437 DEWEKKY-GKNYFENGYDARHKAFLEDSFSLLSSFNKQYETERAVLITQQRWDEYIGELA 495

Qy        463 GITKLGERIKSGKAYADAFEDGKKVEAG----SNITLDAKTGIIDISNSNGKKTQALHFT 518
              |||    |:::  |||||  |  |::||  :|                |   :     |  ||||||    :  :    |  |
Db        496 GITGKGDKLSSGKAYVDYFQEGKLLEKKPDDFSKVVFDPTKGEIDISNS--QTSTLLKFV 553

Qy        519 SPLLTAGTESRERLTNGKYSYINKLKFGRVKNWQVTDG--EASSKLDFSKVIQR------ 570
              :||||  ||||||||      |||  ||   ||                    |  |  :|     :    :    |::  :||
Db        554 TPLLTPGTESRERTQTGKYEYITKLVVKGKDKW-VVNGVKDKGAVYDYTNLIQHAHISSS 612

Qy        571 VAETEGTDEIGLIVNAKAGNDDIFVGQGKMNIDGGDGHDRVFYSKDGGFGNITVDGTSAT 630
               ||   |       |:  |:  :      |||  :|:     |        |    |:|||  |:|  |         |  :  :|||  ||
Db        613 VARGEEYREVRLVSHLGNGNDKVFLAAGSAEIHAGEGHDVVYYDKT-DTGLLVIDGTKAT 671

Qy        631 EAGSYTVNRKVARG-DIYHEVVKRQETKVGKRTETIQYRDYELRKVG-YGYQSTDNLKSV 688
               |  |  |:|  |:::        |     ||:|  |:    ||||  ||::||||||   :   |         ::   |   |  ||
Db        672 EQGRYSVTRELSGATKILREVIKNQKYAVGKREETLEYRDYELTQSGNSNLKAHDELHSV 731

Qy        689 EEVIGSQFNDVFKGSKFNDIFHSGEGDDLLDGGAGDDRLFGGKGNDRLSGD--------- 739
              ||  |||     |  ||||||  ||||        :|||||:|     |||   |:|  ||||  |  ||
Db        732 EE-IGSNQRDEFKGSKFRDIFHGADGDDLLNGNDGDDILYGDKGNDELRGDNGNDQLYGG 790

Qy        740 EGD--------------------------------------------DLLDGGSGDDV 753
              |||                                               ||||||  |:|
Db        791 EGDDKLLGGNGNNYLSGGDGNDELQVLGNGFNVLRGGKGDDKLYGSSGSDLLDGGEGNDY 850

Qy        754 LNGGAGNDVYIFRKGDGNDTLYD---GTGNDKLAFADANISDIMIERTKEGIIVK--RND 808
               |  ||  |:|  |::|      ||  |:||      :  :|||    :|   :    :|:::|      :    :    :        |
Db        851 LEGGDGSDFYVYRSTSGNHTIYDQGKASDSDKLYLSDLSFDNILVKRVNDNLEFRSNNNS 910

Qy        809 HSGSINIPRWYITSNLQNYQSNKTDHKIEQLIGKDGSYITSDQIDKILQDKKDGTVITSQ 868
              :||  :   |    |:      |    |             |||||::  |:|    :|:    :            |         |         |
Db        911 NSGVLTIKDWFKGGNSYN-------HKIEQIVDKNGRKLTAGNLGNNFHD----TQQASS 959

Qy        869 ELKKLADENKSQKLSASDIASSLNKLVGSMALFGTA 904
               ||   :     |       ||:    :   :   |  |::   :        ||   |
Db        960 LLKNVTQEQNESNLSS--LKTELGKIITNAGNFGVA 993
```

RESULT 10

```
HLYA_ECOLI
ID   HLYA_ECOLI     STANDARD;      PRT;  1024 AA.
AC   P08715;
DT   01-JAN-1988 (Rel. 06, Created)
DT   01-JAN-1988 (Rel. 06, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Hemolysin, plasmid.
GN   HLYA.
OS   Escherichia coli.
OG   Plasmid IncI2 pHLY152.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Escherichia.
OX   NCBI_TaxID=562;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Hess J., Wels W., Vogel M., Goebel W.;
RT   "Nucleotide sequence of a plasmid-encoded hemolysin determinant and
RT   its comparison with a corresponding chromosomal hemolysin sequence.";
RL   FEMS Microbiol. Lett. 34:1-11(1986).
RN   [2]
RP   PALMITOYLATION OF LYS-564 AND LYS-690.
RX   MEDLINE=95099325; PubMed=7801126;
RA   Stanley P., Packman L.C., Koronakis V., Hughes C.;
RT   "Fatty acylation of two internal lysine residues required for the
RT   toxic activity of Escherichia coli hemolysin.";
RL   Science 266:1992-1996(1994).
RN   [3]
RP   PALMITOYLATION OF LYS-564 AND LYS-690.
RX   MEDLINE=96404790; PubMed=8808931;
RA   Ludwig A., Garcia F., Bauer S., Jarchau T., Benz R., Hoppe J.,
RA   Goebel W.;
RT   "Analysis of the in vivo activation of hemolysin (HlyA) from
RT   Escherichia coli.";
RL   J. Bacteriol. 178:5422-5430(1996).
CC   -!- FUNCTION: Bacterial hemolysins are exotoxins that attack blood
CC       cell membranes and cause cell rupture by mechanisms not clearly
CC       defined.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- DOMAIN: The Gly-rich region is probably involved in binding
CC       calcium, which is required for target cell-binding or cytolytic
CC       activity.
CC   -!- DOMAIN: The three transmembrane domains are believed to be
CC       involved in pore formation by the cytotoxin.
CC   -!- PTM: Palmitoylated by hlyC. The toxin only becomes active when
CC       modified.
CC   -!- MISCELLANEOUS: The hemolysin of E.Coli is produced predominantly
CC       by strains causing extraintestinal infections, such as those of
CC       the urinary tract.
CC   -!- SIMILARITY: BELONGS TO THE RTX PROKARYOTIC TOXIN FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; M14107; AAA98233.1; -.
DR   InterPro; IPR001343; Hemlysn_Ca_bind.
DR   InterPro; IPR003995; RtxA.
DR   Pfam; PF00353; hemolysinCabind; 6.
DR   Pfam; PF02382; RTX; 1.
DR   PRINTS; PR00313; CABNDNGRPT.
DR   PRINTS; PR01488; RTXTOXINA.
DR   PROSITE; PS00330; HEMOLYSIN_CALCIUM; 4.
KW   Hemolysis; Toxin; Cytolysis; Repeat; Calcium; Transmembrane;
KW   Lipoprotein; Palmitate; Plasmid.
FT   TRANSMEM    238    260       POTENTIAL.
FT   TRANSMEM    268    327       POTENTIAL.
FT   TRANSMEM    365    411       POTENTIAL.
FT   DOMAIN      724    870       16 X REPEATS, GLY-RICH.
FT   REPEAT      724    729       1.
FT   REPEAT      733    738       2.
FT   REPEAT      742    747       3.
FT   REPEAT      751    756       4.
FT   REPEAT      760    765       5.
FT   REPEAT      769    774       6.
FT   REPEAT      778    783       7.
FT   REPEAT      787    792       8.
FT   REPEAT      796    801       9.
FT   REPEAT      807    812       10.
FT   REPEAT      817    822       11.
FT   REPEAT      826    831       12.
FT   REPEAT      835    840       13.
FT   REPEAT      844    849       14.
FT   REPEAT      856    861       15.
FT   REPEAT      865    870       16.
FT   LIPID       564    564       PALMITATE.
FT   LIPID       690    690       PALMITATE.
SQ   SEQUENCE   1024 AA;  110201 MW;  83944917F76C945B CRC64;

  Query Match             39.8%;  Score 1848.5;  DB 1;  Length 1024;
  Best Local Similarity   41.7%;  Pred. No. 1.5e-85;
  Matches  418;  Conservative  156;  Mismatches  312;  Indels  117;  Gaps   22;

Qy          8 KSNIQAGLNSTKSGLKNLYLAIPKDYDPQKGGTLNDFIKAADELGIARLAEEPNHTETAK 67
```

```
                 |  ::      |::     | |  |||||  | |  :|||  ::  |||||     :|  |  |   |
Db         32 KDALKKAAEQTRNAGNRLILLIPKDYKGQ-GSSLNDLVRTADELGIEVQYDEKNGTAITK 90

Qy         68 KSVDTVNQFLSLTQTGIAISATKLEKFLQKH--STNKLAKGLDSVENIDRKLGKASNVLS 125
              :    |  :  : ||: |: |  | :|:| |||:   : |  |     |||   ||||  :||
Db         91 QVFGTAEKLIGLTERGVTIFAPQLDKLLQKYQKAGNILGGG---AENIGDNLGKAGGILS 147

Qy        126 TLSSFLGTALAGIELDSLIKK----GDAAPDALAKASIDLINEIIGNLSQSTQTIEAFSS 181
               |  :||||||: :::| ||||      |: :   ||||||:|||:::  ::      : :||
Db        148 TFQNFLGTALSSMKIDELIKKQKSGGNVSSSELAKASIELINQLVDTVASLNNNVNSFSQ 207

Qy        182 QLAKLGSTISQAKGFSNIGNKLQNL-NFSKTNLGLEIITGLLSGISAGFALADKNASTGK 240
              ||  |||  :|    |    : :|||||||  |      ||: ::|:|| ||| | |:: :| |
Db        208 QLNTLGSVLSNTKHLNGVGNKLQNLPNLDNIGAGLDTVSGILSAISASFILSNADADTRT 267

Qy        241 KVAAGFELSNQVIGNVTKAISSYVLAQRVAAGLSTTGAVAALITSSIMLAISPLAFMNAA 300
               | ||| ||: :|:||| | ||  |::|||  | ||||:  |  | ||  |::  ||||||:|:: |
Db        268 KAAAGVELTTKVLGNVGKGISQYIIAQRAAQGLSTSAAAAGLIASAVTLAISPLSFLSIA 327

Qy        301 DKFNHANALDEFAKQFRKFGYDGDHLLAEYQRGVGTIEASLTTISTALGAVSAGVSAAAV 360
              |||  ||  ::|::::|:| ||||| |||  :  :   | |:|||||||:| |  :||:|:||||
Db        328 DKFKRANKIEEYSQRFKKLGYDGDSLLAAFHKETGAIDASLTTISTVLASVSSGISAAAT 387

Qy        361 GSAVGTPIALLVAGVTGLISGILEASKQAMFESVANRLQGKILEWEKQNGGQNYFDKGYD 420
               | || |:: ||   |||:||||||||||||||||| ||:::    | ||||:: |:|||: |||
Db        388 TSLVGAPVSALVGAVTGIISGILEASKQAMFEHVASKMADVIAEWEKKH-GKNYFENGYD 446

Qy        421 SRYAAYLANNLKFLSELNKELEAERVIAITQQRWDNNIGELAGITKLGERIKSGKAYADA 480
              :|:||:| :| |  ||:  |||    ||  :  ||||  ||   ||||||:|: |::  |||:| |
Db        447 ARHAAFLEDNFKILSQYNKEYSVERSVLITQQHWDTLIGELAGVTRNGDKTLSGKSYIDY 506

Qy        481 FEDGKKVEAGSN----ITLDAKTGIIDISNSNGKKTQALHFTSPLLTAGTESRERLTNGK 536
              :|:||::|     :              |    | ||:|:|    | :   |  |  :||||  |  |  |||   :||
Db        507 YEEGKRLEKKXDEFQKQVFDPLKGNIDLSDS--KSSTLLKFVTPLLTPGEEIRERRQSGK 564

Qy        537 YSYINKLKFGRVKNWQVTD-GEASSKLDFSKVIQRVAETEGT-DEIGLIVNAKAGNDDIF 594
               | ||  :|        |    |  |       :   :    |:| :||  :          ||  :   :       |:| :|
Db        565 YEYITELLVKGVDKWTVKGVQDKGAVYDYSNLIQHASVGNNQYREIRIESHLGDGDDKVF 624

Qy        595 VGQGKMNIDGGDGHDRVFYSK-DGGFGNITVDGTSATEAGSYTVNRKVARGD--IYHEVV 651
              :    |   ||    |  ||| |:|  |  |  |  |:   :|:|||  |||||:|||   |  |    ||   :   |||
Db        625 LSAGSANIYAGKGHDVVYYDKTDTGY--LTIDGTKATEAGNYTVTR-VLGGDVKVLQEVV 681

Qy        652 KRQETKVGKRTETIQYRDYELRKV-GYGYQSTDNLKSVEEVIGSQFNDVFKGSKFNDIFH 710
               | ||  |||||| ||| ||       :  |         |||| ||||:||:     |   | ||||  ||||
Db        682 KEQEVSVGKRTEKTQYRSYEFTHINGKNLTETDNLYSVEELIGTTRADKFFGSKFTDIFH 741

Qy        711 SGEGDDLLDGGAGDDR------------------------------------------- 726
                :||||::|  |:||
Db        742 GADGDDLIEGNDGNDRLYGDKGNDTLSGGNGDDQLYGGDGNDKLIGVAGNNYLNGGDGDD 801

Qy        727 -------------LFGGKGNDRLSGDEGDDLLDGGSGDDVLNGGAGNDVYIFRKGDGNDT 773
                           ||||||||:|  |  ||  ||||||  |||:|  ||  |||:|   :   |  |:
Db        802 EFQVQGNSLAKNVLFGGKGNDKLYGSEGADLLDGGEGDDLLKGGYGNDIYRYLSGYGHHI 861

Qy        774 LYDGTG-NDKLAFADANISDIMIERTKEGIIVKRND-------HSGSINIPRWYITSNLQ 825
              :  |   |  |||: ||  :    |: :|       :|: : :        |       |        |:         :
Db        862 IDDDGGKEDKLSLADIDFRDVAFKREGNDLIMYKGEGNVLSIGHKNGITFRNWF-----E 916

Qy        826 NYQSNKTDHKIEQLIGKDGSYITSDQIDKILQDKK-----------DGTVITSQ------ 868
                 : ::|:|||:    |   |  ||  |  |: ::                |      ||
Db        917 KESGDISNHEIEQIFDKSGRIITPDSLKKALEYQQRNNKASYVYGNDALAYGSQGDLNPL 976

Qy        869 --ELKKLADENKSQKLSASDIASSLNKLVGSMALFGTA-NSVS 908
                |: |:          |   :         |:||  :|  |:  :  |        ||::
Db        977 INEISKIISAAGSFDVKEERTAASLLQLSGNASDFSYGRNSIT 1019
```

```
RESULT 11
HLY1_ECOLI
ID   HLY1_ECOLI     STANDARD;      PRT;  1023 AA.
AC   P09983;
DT   01-MAR-1989 (Rel. 10, Created)
DT   01-MAR-1989 (Rel. 10, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Hemolysin, chromosomal.
GN   HLYA.
OS   Escherichia coli.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Escherichia.
OX   NCBI_TaxID=562;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=J96 / Serotype O4;
RX   MEDLINE=85234404; PubMed=3891743;
RA   Felmlee T., Pellett S., Welch R.A.;
RT   "Nucleotide sequence of an Escherichia coli chromosomal hemolysin.";
RL   J. Bacteriol. 163:94-105(1985).
RN   [2]
RP   SEQUENCE OF 1-44 FROM N.A.
RC   STRAIN=2001;
RX   MEDLINE=85258115; PubMed=3894051;
RA   Nicaud J.-M., Mackman N., Gray L., Holland I.B.;
RT   "Characterisation of HlyC and mechanism of activation and secretion
RT   of haemolysin from E. coli 2001.";
RL   FEBS Lett. 187:339-344(1985).
CC   -!- FUNCTION: Bacterial hemolysins are exotoxins that attack blood
CC       cell membranes and cause cell rupture by mechanisms not clearly
CC       defined.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- DOMAIN: The Gly-rich region is probably involved in binding
CC       calcium, which is required for target cell-binding or cytolytic
CC       activity.
CC   -!- DOMAIN: The three transmembrane domains are believed to be
CC       involved in pore formation by the cytotoxin.
CC   -!- PTM: Palmitoylated by hlyC. The toxin only becomes active when
CC       modified.
CC   -!- MISCELLANEOUS: The hemolysin of E.coli is produced predominantly
CC       by strains causing extraintestinal infections, such as those of
CC       the urinary tract.
CC   -!- SIMILARITY: BELONGS TO THE RTX PROKARYOTIC TOXIN FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; M10133; AAA23975.1; -.
DR   EMBL; X02768; CAA26546.1; -.
DR   PIR; A24433; LEECA.
DR   InterPro; IPR001343; Hemlysn_Ca_bind.
DR   InterPro; IPR003995; RtxA.
DR   Pfam; PF00353; hemolysinCabind; 6.
DR   Pfam; PF02382; RTX; 1.
DR   PRINTS; PR00313; CABNDNGRPT.
DR   PRINTS; PR01488; RTXTOXINA.
DR   PROSITE; PS00330; HEMOLYSIN_CALCIUM; 4.
KW   Hemolysis; Toxin; Cytolysis; Repeat; Calcium; Transmembrane;
KW   Lipoprotein; Palmitate.
FT   TRANSMEM    237    259       POTENTIAL.
FT   TRANSMEM    267    326       POTENTIAL.
FT   TRANSMEM    364    410       POTENTIAL.
FT   DOMAIN      723    869       16 X REPEATS, GLY-RICH.
FT   REPEAT      723    728       1.
FT   REPEAT      732    737       2.
FT   REPEAT      741    746       3.
FT   REPEAT      750    755       4.
FT   REPEAT      759    764       5.
FT   REPEAT      768    773       6.
FT   REPEAT      777    782       7.
FT   REPEAT      786    791       8.
FT   REPEAT      795    800       9.
FT   REPEAT      806    812       10.
```

```
FT   REPEAT      816    821       11.
FT   REPEAT      825    830       12.
FT   REPEAT      834    839       13.
FT   REPEAT      843    848       14.
FT   REPEAT      855    860       15.
FT   REPEAT      864    869       16.
FT   LIPID       563    563       PALMITATE (BY SIMILARITY).
FT   LIPID       689    689       PALMITATE (BY SIMILARITY).
FT   VARIANT       6      6       A -> T (IN STRAIN 2001).
SQ   SEQUENCE   1023 AA;  109867 MW;  196D5C0A9A28B54D CRC64;

  Query Match             39.7%;  Score 1845;  DB 1;  Length 1023;
  Best Local Similarity   41.8%;  Pred. No. 2.2e-85;
  Matches  421;  Conservative  156;  Mismatches  311;  Indels  118;  Gaps   23;

Qy          8 KSNIQAGLNSTKSGLKNLYLAIPKDYDPQKGGTLNDFIKAADELGIARLAEEPNHTETAK 67
              |  ::     |::     | | |||||  | | :||| :: ||||||    :| | |   |
Db         32 KDALKKAAEQTRNAGNRLILLIPKDYKGQ-GSSLNDLVRTADELGIEVQYDEKNGTAITK 90

Qy         68 KSVDTVNQFLSLTQTGIAISATKLEKFLQKH--STNKLAKGLDSVENIDRKLGKASNVLS 125
              :   |  : : ||: |: | | :|:| |||:  : |||     | |||   |||| :|||
Db         91 QVFGTAEKLIGLTERGVTIFAPQLDKLLQKYQKAGNKLG---GSAENIGDNLGKAGSVLS 147

Qy        126 TLSSFLGTALAGIELDSLIKK----GDAAPDALAKASIDLINEIIGNLSQSTQTIEAFSS 181
              |  :||||||: :::| ||||    |: :   ||||||:|||::: : : |    : :||
Db        148 TFQNFLGTALSSMKIDELIKKQKSGGNVSSSELAKASIELINQLV-DTAASLNNVNSFSQ 206

Qy        182 QLAKLGSTISQAKGFSNIGNKLQNL-NFSKTNLGLEIITGLLSGISAGFALADKNASTGK 240
              || |||| :|   |   : :||||||| |       ||: ::|:|| ||| | |:: :| ||
Db        207 QLNKLGSVLSNTKHLNGVGNKLQNLPNLDNIGAGLDTVSGILSAISASFILSNADADTGT 266

Qy        241 KVAAGFELSNQVIGNVTKAISSYVLAQRVAAGLSTTGAVAALITSSIMLAISPLAFMNAA 300
              | ||| ||: :|:||| | || |::||| | ||||: | | || | : ||||||:|:: |
Db        267 KAAAGVELTTKVLGNVGKGISQYIIAQRAAQGLSTSAAAAGLIASVVTLAISPLSFLSIA 326

Qy        301 DKFNHANALDEFAKQFRKFGYDGDHLLAEYQRGVGTIEASLTTISTALGAVSAGVSAAAV 360
              |||  || ::|::::|:| ||||| |||  : :   | |:|||| ||| | :||:|:||||
Db        327 DKFKRANKIEEYSQRFKKLGYDGDSLLAAFHKETGAIDASLTRISTVLASVSSGISAAAT 386

Qy        361 GSAVGTPIALLVAGVTGLISGILEASKQAMFESVANRLQGKILEWEKQNGGQNYFDKGYD 420
               | || |:: ||  |||:|||||||||||||| ||:::   | ||||:: |:|||: |||
Db        387 TSLVGAPVSALVGAVTGIISGILEASKQAMFEHVASKMADVIAEWEKKH-GKNYFENGYD 445

Qy        421 SRYAAYLANNLKFLSELNKELEAERVIAITQQRWDNNIGELAGITKLGERIKSGKAYADA 480
              :|:||:| :| | ||: |||   || : |||| ||  ||||||:|: |::  |||:| |
Db        446 ARHAAFLEDNFKILSQYNKEYSVERSVLITQQHWDTLIGELAGVTRNGDKTLSGKSYIDY 505

Qy        481 FEDGKKVEAG----SNITLDAKTGIIDISNSNGKKTQALHFTSPLLTAGTESRERLTNGK 536
              :|:||::|                 |   | ||:|:|  | :  | | :|||| | | |||   :||
Db        506 YEEGKRLEKKPDEFQKQVFDPLKGNIDLSDS--KSSTLLKFVTPLLTPGEEIRERRQSGK 563

Qy        537 YSYINKLKFGRVKNWQVTD-GEASSKLDFSKVIQRVAETEGT-DEIGLIVNAKAGNDDIF 594
              | || :|    | | |     : |  |:| :||  :        || :  :  |:| :|
Db        564 YEYITELLVKGVDKWTVKGVQDKGSVYDYSNLIQHASVGNNQYREIRIESHLGDGDDKVF 623

Qy        595 VGQGKMNIDGGDGHDRVFYSK-DGGFGNITVDGTSATEAGSYTVNRKVARGD--IYHEVV 651
              :   |  ||  | ||| |:| | | | |:  :|:||| ||||:||| | |  ||  : |||
Db        624 LSAGSANIYAGKGHDVVYYDKTDTGY--LTIDGTKATEAGNYTVTR-VLGGDVKVLQEVV 680

Qy        652 KRQETKVGKRTETIQYRDYELRKV-GYGYQSTDNLKSVEEVIGSQFNDVFKGSKFNDIFH 710
              | ||  |||||| |||| ||    : |      |||| ||||:||:     | | |||| ||||
Db        681 KEQEVSVGKRTEKTQYRSYEFTHINGKNLTETDNLYSVEELIGTTRADKFFGSKFADIFH 740

Qy        711 SGEGDDLLDGGAGDDRLF---------------------------------------- 728
                :||| ::|  |:|||:
Db        741 GADGDDHIEGNDGNDRLYGDKGNDTLSGGNGDDQLYGGDGNDKLIGGAGNNYLNGGDGDD 800

Qy        729 ---------------GGKGNDRLSGDEGDDLLDGGSGDDVLNGGAGNDVYIFRKGDGNDT 773
                              ||||||:| | || |||||| |:|:| || |||:| :  | |:
Db        801 ELQVQGNSLAKNVLSGGKGNDKLYGSEGADLLDGGEGNDLLKGGYGNDIYRYLSGYGHHI 860

Qy        774 LYDGTG-NDKLAFADANISDIMIERTKEGIIVKRND-------HSGSINIPRWYITSNLQ 825
               : |  | :|||: || :  |:   |     :|: : :       |   |   |:    :
Db        861 IDDDGGKDDKLSLADIDFRDVAFRREGNDLIMYKAEGNVLSIGHKNGITFKNWF-----E 915

Qy        826 NYQSNKTDHKIEQLIGKDGSYITSDQIDKILQDKK-----------DGTVITSQ------ 868
                  : ::|:|||:  |||  || | : | |: ::                 |   ||
Db        916 KESGDISNHQIEQIFDKDGRVITPDSLKKALEYQQSNNKASYVYGNDALAYGSQGNLNPL 975

Qy        869 --ELKKLADENKSQKLSASDIASSLNKLVGSMALFGTA-NSVSSNA 911
                |: |:      :   :      |:|| :| |: : |    ||:: |
Db        976 INEISKIISAAGNFDVKEERAAASLLQLSGNASDFSYGRNSITLTA 1021


RESULT 12
RT12_ACTPL
ID   RT12_ACTPL     STANDARD;      PRT;  1023 AA.
AC   P55129;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   RTX-I toxin determinant A from serotypes 5/10 (APX-IA) (Hemolysin IA)
DE   (HLY-IA) (Cytolysin IA) (CLY-IA).
GN   APXIA OR CLYIA OR HLYIA.
OS   Actinobacillus pleuropneumoniae (Haemophilus pleuropneumoniae).
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC   Pasteurellaceae; Actinobacillus.
OX   NCBI_TaxID=715;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=13039 / Serotype 10;
RX   MEDLINE=94276858; PubMed=8007819;
RA   Nagai S., Yagihashi T., Ishihama A.;
RT   "DNA sequence analysis of an allelic variant of the Actinobacillus
RT   pleuropneumoniae-RTX-toxin I (ApxIA) from serotype 10.";
RL   Microb. Pathog. 15:485-495(1993).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K17 / Serotype 5;
RX   MEDLINE=96401417; PubMed=8807793;
RA   Chin N., Frey J., Chang C.F., Chang Y.F.;
RT   "Identification of a locus involved in the utilization of iron by
RT   Actinobacillus pleuropneumoniae.";
RL   FEMS Microbiol. Lett. 143:1-6(1996).
RN   [3]
RP   SEQUENCE OF 886-1023 FROM N.A.
RC   STRAIN=K17 / Serotype 5;
RX   MEDLINE=93366425; PubMed=8359891;
RA   Jansen R., Briaire J., Kamp E.M., Gielkens A.L.J., Smits M.A.;
RT   "Structural analysis of the Actinobacillus pleuropneumoniae-RTX-toxin
RT   I (ApxI) operon.";
RL   Infect. Immun. 61:3688-3695(1993).
CC   -!- FUNCTION: One of the virulence factors of A.pleuropneumoniae,
CC       which has a strong hemolytic activity and is cytotoxic for
CC       alveolar macrophages and neutrophils.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- DOMAIN: The Gly-rich region is probably involved in binding
CC       calcium, which is required for target cell-binding or cytolytic
CC       activity.
CC   -!- DOMAIN: The three transmembrane domains are believed to be
CC       involved in pore formation by the cytotoxin (By similarity).
CC   -!- PTM: PALMITOYLATED BY APXIC. The toxin only becomes active when
CC       modified (By similarity).
CC   -!- MISCELLANEOUS: ApxIA is partially deleted in serotypes 2, 4, 6, 7,
CC       8, 12, and totally deleted in serotype 3.
CC   -!- MISCELLANEOUS: The sequence shown is that of serotype 10.
CC   -!- SIMILARITY: BELONGS TO THE RTX PROKARYOTIC TOXIN FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
```

```
DR   EMBL; D16582; BAA04014.1; -.
DR   EMBL; U04954; AAB17220.1; -.
DR   EMBL; X73116; CAA51546.1; -.
DR   PIR; I39641; I39641.
DR   InterPro; IPR001343; Hemlysn_Ca_bind.
DR   InterPro; IPR003995; RtxA.
DR   Pfam; PF00353; hemolysinCabind; 6.
DR   Pfam; PF02382; RTX; 1.
DR   PRINTS; PR00313; CABNDNGRPT.
DR   PRINTS; PR01488; RTXTOXINA.
DR   PROSITE; PS00330; HEMOLYSIN_CALCIUM; 2.
KW   Hemolysis; Toxin; Cytolysis; Repeat; Calcium; Transmembrane;
KW   Lipoprotein; Palmitate.
FT   TRANSMEM    226    256       POTENTIAL.
FT   TRANSMEM    297    326       POTENTIAL.
FT   TRANSMEM    367    406       POTENTIAL.
FT   DOMAIN      722    845       13 X REPEATS, GLY-RICH.
FT   REPEAT      722    727       1.
FT   REPEAT      731    736       2.
FT   REPEAT      740    745       3.
FT   REPEAT      749    754       4.
FT   REPEAT      758    763       5.
FT   REPEAT      767    772       6.
FT   REPEAT      776    781       7.
FT   REPEAT      785    790       8.
FT   REPEAT      794    799       9.
FT   REPEAT      813    818       10.
FT   REPEAT      822    827       11.
FT   REPEAT      831    836       12.
FT   REPEAT      840    845       13.
FT   CONFLICT    210    217       AMPYLTLA -> GNALSNTR (IN REF. 2).
FT   CONFLICT    581    581       E -> Q (IN REF. 2).
FT   CONFLICT    687    688       TC -> R (IN REF. 2).
FT   CONFLICT   1015   1015       F -> L (IN REF. 2).
SQ   SEQUENCE   1023 AA;  110129 MW;  183C7C15EE57DB55 CRC64;

  Query Match             37.2%;  Score 1729.5;  DB 1;  Length 1023;
  Best Local Similarity   39.5%;  Pred. No. 1.4e-79;
  Matches  403;  Conservative  177;  Mismatches  312;  Indels  127;  Gaps   28;

Qy          8 KSNIQAGLNSTKSGL-------KNLYLAIPKDYDPQKGGTLNDFIKAADELGIARLAEEP 60
              ||  ::|  :  |:||       :  |  |  |||||    |  :|||  :|||:  |||      |
Db         21 KSAAKSGAGALKNGLGQVKQAGQKLILYIPKDYQASTGSSLNDLVKAAEALGIEVHRSEK 80

Qy         61 NHTETAKKSVDTVNQFLSLTQTGIAISATKLEKFLQKHSTNKLAKGL-DSVENIDRKLGK 119
              |  |  ||:     |  :  |   ::  |||:  |  :  :|  |  |:      ||:|  |    |  |  :  ::|  |
Db         81 NGTALAKELFGTTEKLLGFSERGIALFAPQFDKLLNKN--QKLSKSLGGSSEALGQRLNK 138

Qy        120 ASNVLSTLSSFLGTALAGIELDSLIKK----GDAAPDALAKASIDLINEIIGNLSQSTQT 175
                    ||  |  |||||:||::||||:::      |  :     ||||  :||    :::  |:: :|  |
Db        139 TQTALSALQSFLGTAIAGMDLDSLLRRRRNGEDVSGSELAKAGVDLAAQLVDNIASATGT 198

Qy        176 IEAFSSQLAKLGS---TISQAKGFSNIGNKLQNL-NFSKTNLGLEIITGLLSGISAGFAL 231
              ::||:  ||  ||       |::        |  :  :||  ||  :   |       |  :  ::|:||  :||  |   |
Db        199 VDAFAEQLGKLAMPYLTLA----LSGLASKLNNLPDLSLAGPGFDAVSGILSVVSASFIL 254

Qy        232 ADKNASTGKKVAAGFELSNQVIGNVTKAISSYVLAQRVAAGLSTTGAVAALITSSIMLAI 291
              ::|:|    |  |  |||  |:|  :::||:  ||:|  |::||||||||||   |||  |  :  |||
Db        255 SNKDADAGTKAAAGIEISTKILGNIGKAVSQYIIAQRVAAGLSTTAATGGLIGSVVALAI 314

Qy        292 SPLAFMNAADKFNHANALDEFAKQFRKFGYDGDHLLAEYQRGVGTIEASLTTISTALGAV 351
              |||:|:|   ||||     |     |::::::|:||||:||  |||   :   |      |  |||:||||::  |   |
Db        315 SPLSFLNVADKFERAKQLEQYSERFKKFGYEGDSLLASFYRETGAIEAALTTINSVLSAA 374

Qy        352 SAGVSAAAVGSAVGTPIALLVAGVTGLISGILEASKQAMFESVANRLQGKILEWEKQNGG 411
              ||||  |||  ||  ||   |:|  ||:  :||:|||||:|||||:||  ||  :|    ||  ||||::  |
Db        375 SAGVGAAATGSLVGAPVAALVSAITGIISGILDASKQAIFERVATKLANKIDEWEKKH-G 433

Qy        412 QNYFDKGYDSRYAAYLANNLKFLSELNKELEAERVIAITQQRWDNNIGELAGITKLGERI 471
              :|||:  |||:|::|:|   :    :   ||:  |||     |||:||||||||  ||||||||:  |
Db        434 KNYFENGYDARHSAFLEDTFELLSQYNKEYSVERVVAITQQRWDVNIGELAGITRKGADA 493

Qy        472 KSGKAYADAFEDGKKVEAG----SNITLDAKTGIIDISNSNGKKTQALHFTSPLLTAGTE 527
              |||||  |  ||:||  :|                |     |  ||:|:|  |    ||    |  |  :|:  |||  |
Db        494 KSGKAYVDFFEEGKLLEKDPDRFDKKVFDPLEGKIDLSSIN--KTTLLKFITPVFTAGEE 551

Qy        528 SRERLTNGKYSYINKLKFGRVKNWQVTDGEASSKL-DFSKVIQRVAETEGTDEIGLIVNA 586
                |||     |||  |:  :|       :  |  ||    |:  :  :  |::  :||      :  :|  ::     :  :  :
Db        552 IRERKQTGKYEYMTELFVKGKEKWVVTGVESHNAIYDYTNLIQLAIDKKG-EKRQVTIES 610

Qy        587 KAG--NDDIFVGQGKMNIDGGDGHDRVFYSK-DGGFGNITVDGTSATEAGSYTVNRKV-A 642
                  |    ||  |  |::     |          :     |:|||     :|  |     |    |     |:     :|   ||   ||    :||   |    |   :::   |
Db        611 HLGEKNDRIYLSSGSSIVYAGNGHDVAYYDKTDTGY--LTFDGQSAQKAGEYIVTKELKA 668

Qy        643 RGDIYHEVVKRQETKVGKR-TETIQYRDYELR--KVGYGYQSTDNLKSVEEVIGSQFNDV 699
                  :  ·||||  |:    |||   :|   ::||||||        ::|   |   ::   |   |   ||||:||   |
Db        669 DVKVLKEVVKTQDISVGKTCSEKLEYRDYELSPFELGNGIRAKDELHSVEEIIGSNRKDK 728

Qy        700 FKGSKFNDIFHSGEGD-----------------DLLDGGAGDDRLFGGKGNDRLSGDEG 741
                |  ||:|  ||||   :||                    |::  ||  |:|  |  ||  ||||  |  :|
Db        729 FFGSRFTDIFHGAKGDDEIYGNDGHDILYGDDGNDVIHGGDGNDHLVGGNGNDRLIGGKG 788

Qy        742 DDLLDGGSGDD---------VLNGGAGNDVY-------IFRKGDGNDTLYDGTGND--- 781
              :: |:||  |||             ||  ||||||:                 :|   |  |||  :|  |   |   |
Db        789 NNFLNGGDGDDELQVFEGQYNVLLGGAGNDILYGSDGTNLFDGGVGNDKIYGGLGKDIYR 848

Qy        782 ----------------------KLAFADANI----SDIMIERTKEGIIVKRNDHSG-SI 813
                                        |:|  |              |:::  :       |   :          |::|  ::
Db        849 YSKEYGRHIIIEKGGDDDTLLLSDLSFKDVGFIRIGDDLLVNKRIGGTLYYHEDYNGNAL 908

Qy        814 NIPRWYITSNLQNYQSNKTDHKIEQLIGKDGSYITSDQIDKILQDKKDGTVITSQELKKL 873
                    |   |:           :        |:|         |||:::   |||:|:   |     :   ::               :               ||
Db        909 TIKDWF--KEGKEGQNN----KIEKIVDKDGAYVLSQYLTELTAPGRGINYFNGLEEKLY 962

Qy        874 ADENKSQ----KLSASDIASSLNKLVGS----------MALFGTANSVSSNALQPITQ 917
                    |   :            :           |  ||      ||                        :|      :   ||:  :            :||
Db        963 YGEGYNALPQLRKDIEQIISSTGALTGEHGQVLVGAGGPLAYSNSPNSIPNAFSNYLTQ 1021

RESULT 13
RT11_ACTPL
ID   RT11_ACTPL     STANDARD;      PRT;  1023 AA.
AC   P55128;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   RTX-I toxin determinant A from serotypes 1/9 (APX-IA) (Hemolysin IA)
DE   (HLY-IA) (Cytolysin IA) (CLY-IA).
GN   APXIA OR CLYIA OR HLYIA.
OS   Actinobacillus pleuropneumoniae (Haemophilus pleuropneumoniae).
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC   Pasteurellaceae; Actinobacillus.
OX   NCBI_TaxID=715;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=S 4074 / Serotype 1;
RX   MEDLINE=91348845; PubMed=1879928;
RA   Frey J., Meier R., Gygi D., Nicolet J.;
RT   "Nucleotide sequence of the hemolysin I gene from Actinobacillus
RT   pleuropneumoniae.";
RL   Infect. Immun. 59:3026-3032(1991).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=S 4074 / Serotype 1;
RX   MEDLINE=94237497; PubMed=8181764;
RA   Frey J., Haldimann A., Nicolet J., Boffini A., Prentki P.;
RT   "Sequence analysis and transcription of the apxI operon (hemolysin I)
RT   from Actinobacillus pleuropneumoniae.";
RL   Gene 142:97-102(1994).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Isolate CVI 13261 / Serotype 9;
RX   MEDLINE=93366425; PubMed=8359891;
RA   Jansen R., Briaire J., Kamp E.M., Gielkens A.L.J., Smits M.A.;
RT   "Structural analysis of the Actinobacillus pleuropneumoniae-RTX-toxin
RT   I (ApxI) operon.";
```

```
RL   Infect. Immun. 61:3688-3695(1993).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=S 4074 / Serotype 1;
RA   Chang Y., Wang Y., Chin N.;
RL   Submitted (JAN-1994) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: One of the virulence factors of A.pleuropneumoniae,
CC       which has a strong hemolytic activity and is cytotoxic for
CC       alveolar macrophages and neutrophils.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- DOMAIN: The Gly-rich region is probably involved in binding
CC       calcium, which is required for target cell-binding or cytolytic
CC       activity.
CC   -!- DOMAIN: The three transmembrane domains are believed to be
CC       involved in pore formation by the cytotoxin (By similarity).
CC   -!- PTM: Palmitoylated by apxIC. The toxin only becomes active when
CC       modified (By similarity).
CC   -!- MISCELLANEOUS: ApxIA is partially deleted in serotypes 2, 4, 6, 7,
CC       8, 12, and totally deleted in serotype 3.
CC   -!- MISCELLANEOUS: The sequence shown is that of serotype 1.
CC   -!- SIMILARITY: BELONGS TO THE RTX PROKARYOTIC TOXIN FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; X52899; CAA37081.1; -.
DR   EMBL; X68595; CAA48586.1; -.
DR   EMBL; X73117; CAA51548.1; -.
DR   EMBL; U05042; AAB05034.1; -.
DR   PIR; I39643; I39643.
DR   InterPro; IPR001343; Hemlysn_Ca_bind.
DR   InterPro; IPR003995; RtxA.
DR   Pfam; PF00353; hemolysinCabind; 6.
DR   Pfam; PF02382; RTX; 1.
DR   PRINTS; PR00313; CABNDNGRPT.
DR   PRINTS; PR01488; RTXTOXINA.
DR   PROSITE; PS00330; HEMOLYSIN_CALCIUM; 2.
KW   Hemolysis; Toxin; Cytolysis; Repeat; Calcium; Transmembrane;
KW   Lipoprotein; Palmitate.
FT   TRANSMEM    226    256       POTENTIAL.
FT   TRANSMEM    297    326       POTENTIAL.
FT   TRANSMEM    367    406       POTENTIAL.
FT   DOMAIN      722    845       13 X REPEATS, GLY-RICH.
FT   REPEAT      722    727       1.
FT   REPEAT      731    736       2.
FT   REPEAT      740    745       3.
FT   REPEAT      749    754       4.
FT   REPEAT      758    763       5.
FT   REPEAT      767    772       6.
FT   REPEAT      776    781       7.
FT   REPEAT      785    790       8.
FT   REPEAT      794    799       9.
FT   REPEAT      813    818       10.
FT   REPEAT      822    827       11.
FT   REPEAT      831    836       12.
FT   REPEAT      840    845       13.
FT   CONFLICT    210    217       AMPYLTLA -> GNALSNTR (IN REF. 3 AND 4).
FT   CONFLICT    374    374       R -> A (IN REF. 3 AND 4).
FT   CONFLICT    562    562       Q -> E (IN REF. 3 AND 4).
FT   CONFLICT    687    688       TC -> R (IN REF. 3 AND 4).
SQ   SEQUENCE   1023 AA;  110193 MW;  F99A88CFC9F1A598 CRC64;

  Query Match             37.1%;  Score 1723;  DB 1;  Length 1023;
  Best Local Similarity   39.7%;  Pred. No. 3e-79;
  Matches  398;  Conservative  180;  Mismatches  309;  Indels  116;  Gaps   28;

Qy          8 KSNIQAGLNSTKSGL-------KNLYLAIPKDYDPQKGGTLNDFIKAADELGIARLAEEP 60
                ||  ::|  : |:||       : | | |||||    | :||| :|||: |||     |
Db         21 KSAAKSGAGALKNGLGQVKQAGQKLILYIPKDYQASTGSSLNDLVKAAEALGIEVHRSEK 80

Qy         61 NHTETAKKSVDTVNQFLSLTQTGIAIASATKLEKFLQKHSTNKLAKGL-DSVENIDRKLGK 119
              | |  ||:    |  : |   :: |||: | : :| | |:     ||:| |   | | : ::| |
Db         81 NGTALAKELFGTTEKLLGFSERGIALFAPQFDKLLNKN--QKLSKSLGGSSEALGQRLNK 138

Qy        120 ASNVLSTLSSFLGTALAGIELDSLIKK----GDAAPDALAKASIDLINEIIGNLSQSTQT 175
                  || |  |||||:||::||||:::        | :    |||| :||   ::: |:: :| |
Db        139 TQTALSALQSFLGTAIAGMDLDSLLRRRRNGEDVSGSELAKAGVDLAAQLVDNIASATGT 198

Qy        176 IEAFSSQLAKLGS---TISQAKGFSNIGNKLQNL-NFSKTNLGLEIITGLLSGISAGFAL 231
              ::||: || ||       |::       | : :|| || : |     | : ::|:|| :|| | |
Db        199 VDAFAEQLGKLAMPYLTLA----LSGLASKLNNLPDLSLAGPGFDAVSGILSVVSASFIL 254

Qy        232 ADKNASTGKKVAAGFELSNQVIGNVTKAISSYVLAQRVAAGLSTTGAVAALITSSIMLAI 291
              ::|:|  | | ||| |:| :::||: ||:| |::||||||||||  |   || | : |||
Db        255 SNKDADAGTKAAAGIEISTKILGNIGKAVSQYIIAQRVAAGLSTTAATGGLIGSVVALAI 314

Qy        292 SPLAFMNAADKFNHANALDEFAKQFRKFGYDGDHLLAEYQRGVGTIEASLTTISTALGAV 351
              |||:|:| ||||   | |:::::|:||||:|| ||| : |   | |||:||||:: | |
Db        315 SPLSFLNVADKFERAKQLEQYSERFKKFGYEGDSLLASFYRETGAIEAALTTINSVLSAR 374

Qy        352 SAGVSAAAVGSAVGTPIALLVAGVTGLISGILEASKQAMFESVANRLQGKILEWEKQNGG 411
              |||| ||| || || |:| ||: :||:|||||:|||||:|| || :|   || |||:: |
Db        375 SAGVGAAATGSLVGAPVAALVSAITGIISGILDASKQAIFERVATKLANKIDEWEKKH-G 433

Qy        412 QNYFDKGYDSRYAAYLANNLKFLSELNKELEAERVIAITQQRWDNNIGELAGITKLGERI 471
              :|||: |||:|::|:| :  : ||: |||     |||:||||||| ||||||||: |
Db        434 KNYFENGYDARHSAFLEDTFELLSQYNKEYSVERVVAITQQRWDVNIGELAGITRKGSDT 493

Qy        472 KSGKAYADAFEDGKKVEAG----SNITLDAKTGIIDISNSNGKKTQALHFTSPLLTAGTE 527
              |||||| | ||:|| :|                       |  | ||:|: |  ||  |  | :|: ||| |
Db        494 KSGKAYVDFFEEGKLLEKEPDRFDKKVFDPLEGKIDLSSIN--KTTLLKFVTPVFTAGEE 551

Qy        528 SRERLTNGKYSYINKLKFGRVKNWQVTDGEASSKL-DFSKVIQRVAETEGTDEIGLIVNA 586
                |||    ||| |: :|   |    |      : : ::| :||  :: :  : :|  : : : :
Db        552 IRERKQTGKYQYMTELFVKGKEKWVVTGVQSHNAIYDYTNLIQLAIDKKG-EKRQVTIES 610

Qy        587 KAG--NDDIFVGQGKMNIDGGDGHDRVFYSK-DGGFGNITVDGTSATEAGSYTVNRKV-A 642
                | || |::   |    :  |:|||  :|| | | |:  :| || || :|| | | ::: |
Db        611 HLGEKNDRIYLSSGSSIVYAGNGHDVAYYDKTDTGY--LTFDGQSAQKAGEYIVTKELKA 668

Qy        643 RGDIYHEVVKRQETKVGKR-TETIQYRDYELR--KVGYGYQSTDNLKSVEEVIGSQFNDV 699
                :  |||| |:  |||  :| | ::||||||   ::| | :: | | |||:||| |
Db        669 DVKVLKEVVKTQDISVGKTCSEKLEYRDYELSPFELGNGIRAKDELHSVEEIIGSNRKDK 728

Qy        700 FKGSKFNDIFHSGEGD-----------------DLLDGGAGDDRLFGGKGNDRLSGDEG 741
               | ||:| |||| :||                |:: || |:| | || ||||  | :|
Db        729 FFGSRFTDIFHGAKGDDEIYGNDGHDILYGDDGNDVIHGGDGNDHLVGGNGNDRLIGGKG 788

Qy        742 DDLLDGGSGDD----------VLNGGAGNDVY-------IFRKGDGNDTLYDGTGND--- 781
              :: |:|| |||          || ||||||:            :|   | |||  :| | | |
Db        789 NNFLNGGDGDDELQVFEGQYNVLLGGAGNDILYGSDGTNLFDGGVGNDKIYGGLGKDIYR 848

Qy        782 -----------------------KLAFADANI----SDIMIERTKEGIIVKRNDHSG-SI 813
                                     |:| |        |::: :   | :    |::| ::
Db        849 YSKEYGRHIIIEKGGDDDTLLLSDLSFKDVGFIRIGDDLLVNKRIGGTLYYHEDYNGNAL 908

Qy        814 NIPRWYITSNLQNYQSNKTDHKIEQLIGKDGSYITSDQIDKILQDKKDGTVITSQELKKL 873
                | |:     :  |:|    |||::: |||:|: |  : ::     :         | |
Db        909 TIKDWF--KEGKEGQNN----KIEKIVDKDGAYVLSQYLTELTAPGRGINYFNGLEEKLY 962

Qy        874 ADENKSQKLSASDIASSLNKLVGSMALF-GTANSVSSNALQPI 915
                |  :  :  :  : :::  |   | |     ||   :  |:
Db        963 YGEGYN---ALPQLRKDIEQIISSTGAFTGDHGKVSVGSGGPL 1002

RESULT 14
LKTA_ACTAC
ID   LKTA_ACTAC     STANDARD;      PRT;  1050 AA.
AC   P16462;
DT   01-AUG-1990 (Rel. 15, Created)
DT   01-AUG-1990 (Rel. 15, Last sequence update)
```

```
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Leukotoxin.
GN   LKTA OR LTA.
OS   Actinobacillus actinomycetemcomitans (Haemophilus
OS   actinomycetemcomitans).
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC   Pasteurellaceae; Actinobacillus.
OX   NCBI_TaxID=714;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=JP2;
RX   MEDLINE=89359382; PubMed=2670940;
RA   Lally E.T., Golub E.E., Kieba I.R., Taichman N.S., Rosenbloom J.,
RA   Rosenbloom J.C., Gibson C.W., Demuth D.R.;
RT   "Analysis of the Actinobacillus actinomycetemcomitans leukotoxin
RT   gene. Delineation of unique features and comparison to homologous
RT   toxins.";
RL   J. Biol. Chem. 264:15451-15456(1989).
CC   -!- FUNCTION: One of the virulence factors of A.actinomycetemcomitans
CC       might be a cytotoxin, possibly the membrane-bound hemolysin.
CC   -!- SUBCELLULAR LOCATION: Outer-membrane associated or secreted (By
CC       similarity).
CC   -!- DOMAIN: The Gly-rich region is probably involved in binding
CC       calcium, which is required for target cell-binding or cytolytic
CC       activity.
CC   -!- DOMAIN: The three transmembrane domains are believed to be
CC       involved in pore formation by the cytotoxin (BY SIMILARITY).
CC   -!- PTM: Palmitoylated by lktC. The toxin only becomes active when
CC       modified (By similarity).
CC   -!- MISCELLANEOUS: Its target cell specificity is restricted to human
CC       and some non-human cells of the monomyelocytic lineage.
CC   -!- SIMILARITY: BELONGS TO THE RTX PROKARYOTIC TOXIN FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; M27399; AAA21922.1; -.
DR   InterPro; IPR001343; Hemlysn_Ca_bind.
DR   InterPro; IPR003995; RtxA.
DR   Pfam; PF00353; hemolysinCabind; 7.
DR   Pfam; PF02382; RTX; 1.
DR   PRINTS; PR00313; CABNDNGRPT.
DR   PRINTS; PR01488; RTXTOXINA.
DR   PROSITE; PS00330; HEMOLYSIN_CALCIUM; 5.
KW   Hemolysis; Toxin; Cytolysis; Repeat; Calcium; Transmembrane;
KW   Lipoprotein; Palmitate.
FT   TRANSMEM    339    359       POTENTIAL.
FT   TRANSMEM    408    429       POTENTIAL.
FT   TRANSMEM    477    501       POTENTIAL.
FT   DOMAIN      722    844       14 X REPEATS, GLY-RICH.
FT   REPEAT      722    727       1.
FT   REPEAT      731    736       2.
FT   REPEAT      740    745       3.
FT   REPEAT      749    754       4.
FT   REPEAT      758    763       5.
FT   REPEAT      767    772       6.
FT   REPEAT      776    781       7.
FT   REPEAT      785    790       8.
FT   REPEAT      794    799       9.
FT   REPEAT      803    808       10.
FT   REPEAT      812    817       11.
FT   REPEAT      821    826       12.
FT   REPEAT      830    835       13.
FT   REPEAT      839    844       14.
SQ   SEQUENCE   1050 AA;  114194 MW;  38DF9AA24649F662 CRC64;

  Query Match             36.9%;  Score 1712.5;  DB 1;  Length 1050;
  Best Local Similarity   41.2%;  Pred. No. 1e-78;
  Matches  391;  Conservative  167;  Mismatches  292;  Indels   99;  Gaps   21;

Qy          3 NINVIKSNIQAGLNSTKSGLKNLYLAIPKDYDPQKGGTLNDFIKAADELGIARLAEEPNH 62
                ||:  |  :|  |:       | | | |||:|   :||  |   |||| :|||    |  :
Db         27 NIDAAKEQLQKALDKLGKTGKKLTLYIPKNY--KKGNGLTALIKAAQKLGIEVYHEGKDG 84

Qy         63 TETAKKSVDTVNQFLSLTQTGIAISATKLEKFLQ--KHSTNKLAKGLDSVENIDRKLGKA 120
                        ::|  :  | ||: |: :  | :|:|::|  || :|  :      |  |: : : |
Db         85 PALTNGILNTGKKLLGLTERGLTLFAPELDKWIQGNKHLSNSVG----STGNLTKAIDKV 140

Qy        121 SNVLSTLSSFLGTALAGIELDSLIK---KGDAAPDA-LAKASIDLINEIIGNLSQSTQTI 176
                :|| || :|| || :|::||:|||       |      |     |||||::||||:|| :|  |  :
Db        141 QSVLGTLQAFLNTAFSGMDLDALIKARQNGKNVTDVQLAKASLNLINELIGTISSITNNV 200

Qy        177 EAFSSQLAKLGSTISQAKGFSNIGNKLQNL----NFSKTNLGLEIITGLLSGISAGFALA 232
                : || || |||   :  |  |  |  :  |:||:||     |     |     ::|:|| |||  ||
Db        201 DTFSKQLNKLGEALGQVKHFGSFGDKLKNLPKLGNLGK---GYGALSGVLSAISAALLLA 257

Qy        233 DKNASTGKKVAAGFELSNQVIGNVTKAISSYVLAQRVAAGLSTTGAVAALITSSIMLAIS 292
                :|:|  |    | ||   ||:|:|:||: |||: |::||| |||||||| || || |  : ||||
Db        258 NKHADTATKAAAAAELTNKVLGNIGKAITQYLIAQRAAAGLSTTGPVAGLIASVVSLAIS 317

Qy        293 PLAFMNAADKFNHANALDEFAKQFRKFGYDGDHLLAEYQRGVGTIEASLTTISTALGAVS 352
                ||:|:  |  :|: |  |: ::|:|:|:||||:|| || :: :   |   :|::|||:| | |::
Db        318 PLSFLGIAKQFDRARMLEAYSKRFKKFGYNGDSLLGQFYKNTGIADAAITTINTVLSAIA 377

Qy        353 AGVSAAAVGSAVGTPIALLVAGVTGLISGILEASKQAMFESVANRLQGKILEWEKQNGGQ 412
                ||| ||: || || || |||: :| ||||||:|||||: | :||:|  ||  || :  |:
Db        378 AGVGAASAGSLVGAPIGLLVSAITSLISGILDASKQAVSEHIANQLADKIKAWENKY-GK 436

Qy        413 NYFDKGYDSRYAAYLANNLKFLSELNKELEAERVIAITQQRWDNNIGELAGITKLGERIK 472
                || : |||:|::|:| ::||  :|| :: : | :::|||| || |||||||:|:||:
Db        437 NYSENGYDARHSAFLEDSLKLFNELREKYKTENILSITQQGWDQRIGELAGITRNGDRIQ 496

Qy        473 SGKAYADAFEDGKKVEAGSN----ITLDAKTGIIDISNSNGKKTQALHFTSPLLTAGTES 528
                ||||| |  : |:::   |:       ||   | ||:|   |   |   |  | :||||||  |
Db        497 SGKAYVDYLKKGEELAKHSDKFTKQILDPIKGNIDLSGIKGSTT--LTFLNPLLTAGKEE 554

Qy        529 RERLTNGKYSYINKLKFGRVKNWQVTDGEASSKL-DFSKVIQR-VAETEGTDEIGLIVNA 586
                |:   :||| :| :||      :|:|      |: : ||| :|| |          |  |||
Db        555 RKTRQSGKYEFITELKVKGRTDWKVKGVPNSNGVYDFSNLIQHAVTRDNKVLEARLIANL 614

Qy        587 KAGNDDIFVGQGKMNIDGGDGHDRVFYSKDGGFGNITVDGTSATEAGSYTVNRKVARGDI 646
                | :| :||| |   :: |||:| | ||| |  | |:|:|| :||:|| | | | ::   :
Db        615 GAKDDYVFVGSGSTIVNAGDGYDVVDYSK-GRTGALTIDGRNATKAGQYKVERDLSGTQV 673

Qy        647 YHEVVKRQETKVGKRTETIQYRDYEL--RKVGYGYQSTDNLKSVEEVIGSQFNDVFKGSK 704
                  |  | :|||| || |: ::|:|:||          |:::  |  | ||||:|||  |  |||
Db        674 LQETVSKQETKRGKVTDLLEYRNYKLDYYYTNKGFKAHDELNSVEEIIGSNLRDKFYGSK 733

Qy        705 FNDIFHSGEGDDL---------------------------LDGGAGDDRLF--------- 728
                |||:||  :||||                           | ||||:||||
Db        734 FNDVFHGHDGDDLIYGYDGDDRLYGDNGNDEIHGGQGNDKLYGGAGNDRLFGEYGNNYLD 793

Qy        729 -----------------GGKGNDRLSGDEGDDLLDGGSGDDVLNGGAGNDVYIFRKGDG 770
                              || ||:| |::||||||| ||| | || |||:|::|| |
Db        794 GGEGDDHLEGGNGSDILRGGSGNDKLFGNQGDDLLDGGEGDDQLAGGEGNDIYVYRKEYG 853

Qy        771 NDTLYDGTGN-DKLAFADANISDIMIERTKEGIIVKRNDHSGSINIPRWYITSN----LQ 825
                : |: : :|: |||: |: |: |: ||   :::| |: : ::   |:   |    |
Db        854 HHTITEHSGDKDKLSLANINLKDVSFERNGNDLLLKTNNRT-AVTFKGWFSKPNSSAGLD 912

Qy        826 NYQSNKTDHKIEQLIGKDGSYITSDQIDKILQDKKDGTVITSQELKKLA 874
                ||     ::   |    ||      |   :       :   :||||:
Db        913 EYQRKLLEYAPE----KD----VHDLRDNLSYSEVKSTNHSIIKLKKLS 953


RESULT 15
CYAA_BORPE
ID   CYAA_BORPE     STANDARD;      PRT;  1706 AA.
AC   P15318;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-APR-1990 (Rel. 14, Last sequence update)
```

```
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Bifunctional hemolysin-adenylate cyclase precursor (Cyclolysin) (ACT)
DE   (AC-HLY) [Contains: Calmodulin-sensitive adenylate cyclase
DE   (EC 4.6.1.1) (ATP pyrophosphate-lyase) (Adenylyl cyclase); Hemolysin].
GN   CYA OR CYAA.
OS   Bordetella pertussis.
OC   Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC   Alcaligenaceae; Bordetella.
OX   NCBI_TaxID=520;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=18323;
RX   MEDLINE=88216178; PubMed=2897067;
RA   Glaser P., Ladant D., Sezer O., Pichot F., Ullmann A., Danchin A.;
RT   "The calmodulin-sensitive adenylate cyclase of Bordetella pertussis:
RT   cloning and expression in Escherichia coli.";
RL   Mol. Microbiol. 2:19-30(1988).
RN   [2]
RP   SEQUENCE OF 1489-1706 FROM N.A., AND BIFUNCTIONAL PROTEIN DESCRIPTION.
RC   STRAIN=18323;
RX   MEDLINE=89091151; PubMed=2905265;
RA   Glaser P., Sakamoto H., Bellalou J., Ullmann A., Danchin A.;
RT   "Secretion of cyclolysin, the calmodulin-sensitive adenylate cyclase-
RT   haemolysin bifunctional protein of Bordetella pertussis.";
RL   EMBO J. 7:3997-4004(1988).
RN   [3]
RP   DOMAINS.
RX   MEDLINE=91177021; PubMed=2007407;
RA   Munier H., Gilles A.-M., Glaser P., Danchin A., Sarfati R., Barzu O.;
RT   "Isolation and characterization of catalytic and calmodulin-binding
RT   domains of Bordetella pertussis adenylate cyclase.";
RL   Eur. J. Biochem. 196:469-474(1991).
RN   [4]
RP   MUTAGENESIS.
RX   MEDLINE=89251630; PubMed=2542030;
RA   Glaser P., Elmaoglou-Lazaridou A., Krin E., Ladant D., Barzu O.,
RA   Danchin A.;
RT   "Identification of residues essential for catalysis and binding of
RT   calmodulin in Bordetella pertussis adenylate cyclase by site-directed
RT   mutagenesis.";
RL   EMBO J. 8:967-972(1989).
RN   [5]
RP   MUTAGENESIS.
RX   MEDLINE=91266896; PubMed=2050107;
RA   Glaser P., Munier H., Gilles A.-M., Krin E., Porumb T., Barzu O.,
RA   Sarfati R., Pellecuer C., Danchin A.;
RT   "Functional consequences of single amino acid substitutions in
RT   calmodulin-activated adenylate cyclase of Bordetella pertussis.";
RL   EMBO J. 10:1683-1688(1991).
RN   [6]
RP   REVIEW.
RX   MEDLINE=93119764; PubMed=8418825;
RA   Danchin A.;
RT   "Phylogeny of adenylyl cyclases.";
RL   Adv. Second Messenger Phosphoprotein Res. 27:109-162(1993).
RN   [7]
RP   PALMITOYLATION OF LYS-983.
RX   MEDLINE=95025937; PubMed=7939682;
RA   Hackett M., Guo L., Shabanowitz J., Hunt D.F., Hewlett E.L.;
RT   "Internal lysine palmitoylation in adenylate cyclase toxin from
RT   Bordetella pertussis.";
RL   Science 266:433-435(1994).
RN   [8]
RP   PALMITOYLATION OF LYS-860.
RX   MEDLINE=99214144; PubMed=10196151;
RA   Basar T., Havlicek V., Bezouskova S., Halada P., Hackett M., Sebo P.;
RT   "The conserved lysine 860 in the additional fatty-acylation site of
RT   Bordetella pertussis adenylate cyclase is crucial for toxin function
RT   independently of its acylation status.";
RL   J. Biol. Chem. 274:10777-10783(1999).
CC   -!- FUNCTION: THIS ADENYLATE CYCLASE BELONGS TO A SPECIAL CLASS OF
CC       BACTERIAL TOXIN. IT CAUSES WHOOPING COUGH BY ACTING ON MAMMALIAN
CC       CELLS BY ELEVATING CAMP-CONCENTRATION AND THUS DISRUPTS NORMAL
CC       CELL FUNCTION.
CC   -!- CATALYTIC ACTIVITY: ATP = 3',5'-cyclic AMP + diphosphate.
CC   -!- ENZYME REGULATION: ACTIVATED BY HOST CALMODULIN.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- DOMAIN: THE GLY-RICH REGION IS PROBABLY INVOLVED IN BINDING
CC       CALCIUM, WHICH IS REQUIRED FOR TARGET CELL-BINDING OR CYTOLYTIC
CC       ACTIVITY (BY SIMILARITY).
CC   -!- PTM: RELEASED IN A PROCESSED FORM.
CC   -!- PTM: PALMITOYLATED BY CYAC. THE TOXIN ONLY BECOMES ACTIVE WHEN
CC       MODIFIED IN POSITION LYS-983.
CC   -!- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE ADENYLYL
CC       CYCLASE CLASS-2 FAMILY.
CC   -!- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE RTX
CC       PROKARYOTIC TOXIN FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; Y00545; CAA68613.1; -.
DR   EMBL; X14199; CAA32411.1; -.
DR   EMBL; A07292; CAA00653.1; -.
DR   EMBL; A14850; CAA01202.1; -.
DR   PIR; S00893; OYBRC.
DR   HSSP; P40136; 1K90.
DR   InterPro; IPR005165; Anthrax_toxA.
DR   InterPro; IPR001343; Hemlysn_Ca_bind.
DR   InterPro; IPR003995; RtxA.
DR   Pfam; PF03497; Anthrax_toxA; 1.
DR   Pfam; PF00353; hemolysinCabind; 17.
DR   Pfam; PF02382; RTX; 1.
DR   PRINTS; PR00313; CABNDNGRPT.
DR   PRINTS; PR01488; RTXTOXINA.
DR   PROSITE; PS00330; HEMOLYSIN_CALCIUM; 5.
KW   Lyase; cAMP biosynthesis; ATP-binding; Hemolysis; Toxin; Virulence;
KW   Whooping cough; Calcium-binding; Repeat; Lipoprotein; Palmitate.
FT   CHAIN         1    312       CALMODULIN-SENSITIVE ADENYLATE
FT                                CYCLASE.
FT   CHAIN       313   1706       HEMOLYSIN (BY SIMILARITY TO E.COLI
FT                                HEMOLYSIN HLYA).
FT   DOMAIN        1    399       A, CATALYTIC.
FT   DOMAIN      400    912       B, ALA/GLY-RICH.
FT   DOMAIN      913   1656       C.
FT   DOMAIN     1657   1706       D, ASP/GLY-RICH.
FT   NP_BIND     349    356       ATP (POTENTIAL).
FT   DOMAIN      913   1610       28 X REPEATS, GLY-RICH.
FT   REPEAT      913    918       1.
FT   REPEAT     1015   1020       2.
FT   REPEAT     1024   1029       3.
FT   REPEAT     1033   1038       3.
FT   REPEAT     1042   1047       4.
FT   REPEAT     1051   1056       5.
FT   REPEAT     1060   1065       6.
FT   REPEAT     1080   1085       7.
FT   REPEAT     1165   1170       8.
FT   REPEAT     1174   1179       9.
FT   REPEAT     1183   1188       10.
FT   REPEAT     1203   1208       11.
FT   REPEAT     1280   1285       12.
FT   REPEAT     1289   1294       13.
FT   REPEAT     1298   1303       14.
FT   REPEAT     1307   1312       15.
FT   REPEAT     1316   1321       16.
FT   REPEAT     1325   1330       17.
FT   REPEAT     1345   1350       18.
FT   REPEAT     1421   1426       19.
FT   REPEAT     1430   1435       20.
FT   REPEAT     1439   1444       21.
FT   REPEAT     1448   1453       22.
```

```
FT   REPEAT      1556   1561       23.
FT   REPEAT      1565   1570       24.
FT   REPEAT      1574   1579       25.
FT   REPEAT      1583   1588       26.
FT   REPEAT      1593   1598       27.
FT   REPEAT      1605   1610       28.
FT   LIPID        860    860       PALMITATE.
FT   LIPID        983    983       PALMITATE.
FT   MUTAGEN      188    188       D->E,N,Y,H: LOSS OF ACTIVITY.
FT   MUTAGEN      190    190       D->N,Y,H: LOSS OF ACTIVITY.
FT   MUTAGEN      298    298       H->R,P,L: LOSS OF ACTIVITY.
FT   MUTAGEN      301    301       E->Q,K: LOSS OF ACTIVITY.
SQ   SEQUENCE   1706 AA;  177506 MW;  F00744524BDD442E CRC64;

  Query Match             17.3%;  Score 802;  DB 1;  Length 1706;
  Best Local Similarity   27.1%;  Pred. No. 1.1e-32;
  Matches  246;  Conservative  142;  Mismatches  347;  Indels  174;  Gaps   26;

Qy        106 GLDSVENIDRKLGKASNVLSTLSSFLGTALAGIELDSLI-----KKGDAAP--------- 151
              | ||:: :      |   |   |   :|       | :|:   :      :: || |
Db        403 GYDSLDGV----GSRSFSLGEVSDMAAVEAAELEMTRQVLHAGARQDDAEPGVSGASAHW 458

Qy        152 --------DALAKA-----SIDLINEI----IGNLSQSTQTIEAFSSQLAKLGSTISQ-A 193
                      |:| |      :| |: :        |   |    :: |     | :   | :::
Db        459 GQRALQGAQAVAAAQRLVHAIALMTQFGRAGSTNTPQEAASLSAAVFGLGEASSAVAETV 518

Qy        194 KGFSNIGNKLQNLNFSKTNLGLEIITGLLSGISAGFALADKNASTGKKVAAGFELSNQVI 253
               ||    |:         |          : :   |: : : || :| | :|   |:| ||| |:: |:
Db        519 SGFFR-GSSRWAGGFGVAGGAMALGGGIAAAVGAGMSLTD-DAPAGQKAAAGAEIALQLT 576

Qy        254 GNVTKAISSYVL----AQRVAAGLSTTGAVAALITSSIMLAISPLAFMNAADKFNHANAL 309
              |     :  ||  |      |: | :||    |||         ::  |:||:          : ::|: |
Db        577 GGTVELASSIALALAAARGVTSGLQVAGASAGAAAGALAAALSPMEIYGLVQQSHYADQL 636

Qy        310 DEFAKQFRKFGYDGDHLLAEYQRGVGTIEASLTTISTALGAVSAGVSAAAVGSAVGTPIA 369
              |: |::    :||:|| |||:   |          | ::   :|    |    |   | |  ||   | || |:|
Db        637 DKLAQESSAYGYEGDALLAQLYRDKTAAEGAVAGVSAVLSTVGAAVSIAAAASVVGAPVA 696

Qy        370 LLVAGVTGLISGILEASKQAMFESVANRLQGKILEWEKQNGGQNYFDKGYDSRYAAYLAN 429
              :: : :|| ::|||    :| :  | :||        ||     ::    |  | ||:|     :|:    |||
Db        697 VVTSLLTGALNGILRGVQQPIIEKLANDYARKI---DELGGPQAYFEKNLQARH-EQLAN 752

Qy        430 N---LKFLSELNKELEAERVIAITQQRWDNNIGELAGITKLGERIKSGKAYADAFEDGKK 486
               :       | |::|       |  || :          :    ||| ||     : :||     :  |  |    |::
Db        753 SDGLRKMLADLQAGWNASSVIGVQTTEISKSALELAAITGNADNLKSVDVFVDRFVQGER 812

Qy        487 VEAGSNITLDAKTGIIDISNSNGKKTQALHFTSPLLTAGTESRERLTNGKYSYINKLKF- 545
              | ||    :  ||   | |||:   |::   || :|| |      |   |   ||      :   ::
Db        813 V-AGQPVVLDVAAGGIDIASRKGER-PALTFITPLAAPGEEQRRRTKTGKSEFTTFVEIV 870

Qy        546 GRVKNWQVTDGEASSKLDFSKVIQRVAETEGT--DEIGLIVNAKAGNDDIFVGQGKMNID 603
              |:    |::  || | :  :|  :||: :: :     |    |       |:|  :       :::  |
Db        871 GKQDRWRIRDGAADTTIDLAKVVSQLVDANGVLKHSIKLDVIGGDGDDVVLANASRIHYD 930

Qy        604 GGDGHDRVFYSKDGGFGNITVDGTSATEAGSYTVNRKVARGDIYHEVVKRQETKVGKRTE 663
              || |  : |  |:    |    :|||      :  :    :  | :::     ::| |  |   |  |   |||||
Db        931 GGAGTNTVSYAALGRQDSITV----SADGERFNVRKQLNNANVYREGVATQTTAYGKRTE 986

Qy        664 TIQYRDYELRKVGYGYQSTDNLKSVEEVIGSQFNDVFKGSKFNDIFHSGEGDDLLDGGAG 723
               :|||  || :||      :    |  |: |: :||    ||     |:   ::        | ||| |||||
Db        987 NVQYRHVELARVGQVVE-VDTLEHVQHIIGGAGNDSITGNAHDNFLAGGSGDDRLDGGAG 1045

Qy        724 DDRLFGGKG--------------------------------------------------- 732
              :| |  ||:|
Db       1046 NDTLVGGEGQNTVIGGAGDDVFLQDLGVWSNQLDGGAGVDTVKYNVHQPSEERLERMGDT 1105

Qy        733 -----------------------------------NDRLSGDEGDDLLDGGSGDDVLNG 756
                                                 |||::||:  |:  |  |   |:|  :  |
Db       1106 GIHADLQKGTVEKWPALNLFSVDHVKNIENLHGSRLNDRIAGDDQDNELWGHDGNDTIRG 1165

Qy        757 GAGNDVYIFRKGDGNDTLYDGTGNDKLAFADANISDIM-----IERTKEGIIVKRNDHSG 811
                  |:|   |  |  |  | ||||    |||        :||  :        ::        ::      |  |
Db       1166 RGGDD--ILRGGLGLDTLYGEDGNDIFLQDDETVSDDIDGGAGLDTVDYSAMI----HPG 1219

Qy        812 SINIPRWY---ITSNLQNYQSNKTD----------HKIEQLIG---KDGSYITSDQIDKI 855
                   |  |   |   |  . | ::|        |                   :| :||   ||   |    |  : :
Db       1220 RIVAPHEYGFGIEADLSREWVRKASALGVDYYDNVRNVENVIGTSMKD-VLIGDAQANTL 1278

Qy        856 LQDKKDGTV 864
               :    | ||
Db       1279 MGQGGDDTV 1287


RESULT 16
CYAA_BORBR
ID   CYAA_BORBR     STANDARD;      PRT;  1705 AA.
AC   Q57506; O05179;
DT   01-NOV-1997 (Rel. 35, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Bifunctional hemolysin-adenylate cyclase precursor (Cyclolysin) (ACT)
DE   (AC-HLY) [Contains: Calmodulin-sensitive adenylate cyclase
DE   (EC 4.6.1.1) (ATP pyrophosphate-lyase) (Adenylyl cyclase); Hemolysin].
GN   CYA OR CYAA.
OS   Bordetella bronchiseptica (Alcaligenes bronchisepticus).
OC   Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC   Alcaligenaceae; Bordetella.
OX   NCBI_TaxID=518;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CIP 9.73;
RX   MEDLINE=96009899; PubMed=7557410;
RA   Betsou F., Sismeiro O., Danchin A., Guiso N.;
RT   "Cloning and sequence of the Bordetella bronchiseptica adenylate
RT   cyclase-hemolysin-encoding gene: comparison with the Bordetella
RT   pertussis gene.";
RL   Gene 162:165-166(1995).
RN   [2]
RP   REVISIONS TO 1555-1558.
RA   Danchin A., Boursaux-Eude C.;
RL   Submitted (DEC-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: THIS ADENYLATE CYCLASE BELONGS TO A SPECIAL CLASS OF
CC       BACTERIAL TOXIN. IT CAUSES WHOOPING COUGH BY ACTING ON MAMMALIAN
CC       CELLS BY ELEVATING CAMP-CONCENTRATION AND THUS DISRUPTS NORMAL
CC       CELL FUNCTION.
CC   -!- CATALYTIC ACTIVITY: ATP = 3',5'-cyclic AMP + diphosphate.
CC   -!- ENZYME REGULATION: ACTIVATED BY HOST CALMODULIN.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- DOMAIN: THE GLY-RICH REGION IS PROBABLY INVOLVED IN BINDING
CC       CALCIUM, WHICH IS REQUIRED FOR TARGET CELL-BINDING OR CYTOLYTIC
CC       ACTIVITY (BY SIMILARITY).
CC   -!- PTM: RELEASED IN A PROCESSED FORM.
CC   -!- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE ADENYLYL
CC       CYCLASE CLASS-2 FAMILY.
CC   -!- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE RTX
CC       PROKARYOTIC TOXIN FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; Z37112; CAA85481.2; -.
DR   HSSP; P40136; 1K90.
DR   InterPro; IPR005165; Anthrax_toxA.
DR   InterPro; IPR001343; Hemlysn_Ca_bind.
DR   InterPro; IPR003995; RtxA.
DR   Pfam; PF03497; Anthrax_toxA; 1.
DR   Pfam; PF00353; hemolysinCabind; 17.
DR   Pfam; PF02382; RTX; 1.
DR   PRINTS; PR00313; CABNDNGRPT.
DR   PRINTS; PR01488; RTXTOXINA.
DR   PROSITE; PS00330; HEMOLYSIN_CALCIUM; 5.
```

```
KW   Lyase; cAMP biosynthesis; ATP-binding; Hemolysis; Toxin; Virulence;
KW   Whooping cough; Calcium-binding; Repeat; Lipoprotein; Palmitate.
FT   CHAIN         1    312       CALMODULIN-SENSITIVE ADENYLATE
FT                                CYCLASE.
FT   CHAIN       313   1705       HEMOLYSIN (BY SIMILARITY TO E.COLI
FT                                HEMOLYSIN HLYA).
FT   DOMAIN        1    399       A, CATALYTIC.
FT   DOMAIN      400    911       B, ALA/GLY-RICH.
FT   DOMAIN      912   1655       C.
FT   DOMAIN     1656   1705       D, ASP/GLY-RICH.
FT   NP_BIND     349    356       ATP (POTENTIAL).
FT   DOMAIN      912   1609       28 X 5 AA REPEATS, GLY-RICH.
FT   REPEAT      912    917       1.
FT   REPEAT     1014   1019       2.
FT   REPEAT     1023   1028       3.
FT   REPEAT     1032   1037       3.
FT   REPEAT     1041   1046       4.
FT   REPEAT     1050   1055       5.
FT   REPEAT     1059   1064       6.
FT   REPEAT     1079   1084       7.
FT   REPEAT     1164   1169       8.
FT   REPEAT     1173   1178       9.
FT   REPEAT     1182   1187       10.
FT   REPEAT     1202   1207       11.
FT   REPEAT     1279   1284       12.
FT   REPEAT     1288   1293       13.
FT   REPEAT     1297   1302       14.
FT   REPEAT     1306   1311       15.
FT   REPEAT     1315   1320       16.
FT   REPEAT     1324   1329       17.
FT   REPEAT     1344   1349       18.
FT   REPEAT     1420   1425       19.
FT   REPEAT     1429   1434       20.
FT   REPEAT     1438   1443       21.
FT   REPEAT     1447   1452       22.
FT   REPEAT     1555   1560       23.
FT   REPEAT     1564   1569       24.
FT   REPEAT     1573   1578       25.
FT   REPEAT     1582   1587       26.
FT   REPEAT     1592   1597       27.
FT   REPEAT     1604   1609       28.
FT   LIPID       859    859       PALMITATE (BY SIMILARITY).
FT   LIPID       982    982       PALMITATE (BY SIMILARITY).
SQ   SEQUENCE   1705 AA;  177153 MW;  C43B30F5886C835A CRC64;

  Query Match             17.2%;  Score 801.5;  DB 1;  Length 1705;
  Best Local Similarity   27.7%;  Pred. No. 1.1e-32;
  Matches  254;  Conservative  135;  Mismatches  336;  Indels  193;  Gaps   28;

Qy        106 GLDSVENIDRKLGKASNVLSTLSSFLGTALAGIELDSLI-----KKGDAAP--------- 151
              | ||:: :      |   |    |   :|          | :|:   :        :: || |
Db        403 GYDSLDGV----GSRSFSLGEVSDMAAVEAAELEMTRQVLHAGARQDDAEPGVSGASAHW 458

Qy        152 ------DALAKASIDLINEIIGNLSQ-----STQT-IEAFSSQLAKLG---------STI 190
                    | | |:    :     | ::|       || |  ||  |    |    |        |:
Db        459 GQRALQGAQAVAAAQRLVHAIALMTQFGRAGSTNTPQEAASLSAAVFGLGEASSAVAETV 518

Qy        191 S--------QAKGFSNIGNKLQNLNFSKTNLGLEIITGLLSGISAGFALADKNASTGKKV 242
              |               | ||   |                  |   : | :  : || :| | :|  |:|
Db        519 SGFFRGSSRWAGGFGVAG-------------GAMALGGGIGAVGAGMSLTD-DAPAGQKA 564

Qy        243 AAGFELSNQVIGNVTKAISSYVL----AQRVAAGLSTTGAVAALITSSIMLAISPLAFMN 298
              ||| |:: |: |    :   ||   |      |: | :||   || |       :: |:||:
Db        565 AAGAEIALQLTGGTVELASSIALALAAARGVTSGLQVAGASAGAAAGALAAALSPMEIYG 624

Qy        299 AADKFNHANALDEFAKQFRKFGYDGDHLLAEYQRGVGTIEASLTTISTALGAVSAGVSAA 358
                : ::|: ||: |::    :||:|| |||:   |        | ::   :|   |   | | || |
Db        625 LVQQSHYADQLDKLAQESSAYGYEGDALLAQLYRDKTAAEGAVAGVSAVLSTVGAAVSIA 684

Qy        359 AVGSAVGTPIALLVAGVTGLISGILEASKQAMFESVANRLQGKILEWEKQNGGQNYFDKG 418
              |  | || |:|:: : :|| :|||        :|  :  : :||      ||    :: | ||:|
Db        685 AAASVVGAPVAVVTSLLTGALNGILRGVQQPIIEKLANDYARKI---DELGGPQAYFEKN 741

Qy        419 YDSRYAAYLANN---LKFLSELNKELEAERVIAITQQRWDNNIGELAGITKLGERIKSGK 475
                 :|:   |||:    | |::|    |   || :         :  ||| ||    : :||
Db        742 LQARH-EQLANSDGLRKMLADLQAGWNASSVIGVQTTEISKSALELAAITGNADNLKSAD 800

Qy        476 AYADAFEDGKKVEAGSNITLDAKTGIIDISNSNGKKTQALHFTSPLLTAGTESRERLTNG 535
                : | |  |::| ||  :  ||   | |||::  |::   || | :||   | | | |    |
Db        801 VFVDRFIQGERV-AGQPVVLDVAAGGIDIASRKGER-PALTFITPLAAPGEEQRRRTKTG 858

Qy        536 KYSYINKLKF-GRVKNWQVTDGEASSKLDFSKVIQRVAETEGT--DEIGLIVNAKAGNDD 592
              |  :   ::  |:      |:: || |  : :| :||: :: :    |         | | |     |:|
Db        859 KSEFTTFVEIVGKQDRWRIRDGAADTTIDLAKVVSQLVDANGVLKHSIKLEVIGGDGDDV 918

Qy        593 IFVGQGKMNIDGGDGHDRVFYSKDGGFGNITVDGTSATEAGSYTVNRKVARGDIYHEVVK 652
              :       ::: ||| | :  | |:   |     :|||      : :    : | :::    ::| | |
Db        919 VLANASRIHYDGGAGTNTVSYAALGRQDSITV----SADGERFNVRKQLNNANVYREGVA 974

Qy        653 RQETKVGKRTETIQYRDYELRKVGYGYQSTDNLKSVEEVIGSQFNDVFKGSKFNDIFHSG 712
              |:|   ||||| :|||   || :||             | |: |: :||    ||    |:  ::    |
Db        975 TQKTAYGKRTENVQYRHVELARVGQ-LVEVDTLEHVQHIIGGAGNDSITGNAHDNFLAGG 1033

Qy        713 EGDDLLDGGAGDDRLFGGKGNDRLSGDEGDDL-----------LDGGSG----------- 750
              ||| |||||:| | ||:|:: :  |    |||:                 ||||:|
Db       1034 AGDDRLDGGAGNDTLVGGEGHNTVVGGAGDDVFLQDLGVWSNQLDGGAGVDTVKYNVHQP 1093

Qy        751 ---------------------------------------------------------DDVL 754
                                                                       |: |
Db       1094 SEERLERMGDTGIHADLQKGTVEKWPALNLFSVDHVKNIENLHGSSLNDSIAGDDRDNEL 1153

Qy        755 NGGAGNDVY-------IFRKGDGNDTLYDGTGNDKLAFADANISDIM-----IERTKEGI 802
               |  |||          | | | |  ||||   |||       |  :|| :          ::
Db       1154 WGDDGNDTIHGRGGDDILRGGLGLDTLYGEDGNDIFLQDDETVSDDIDGGAGLDTVDYSA 1213

Qy        803 IVKRNDHSGSINIPRWY---ITSNLQNYQSNKTDHK----------IEQLIG---KDGSY 846
              ::       |:| |   |   |    |  ::|          |    :          :| :||   ||
Db       1214 MI----HAGKIVAPHEYGFGIEADLSEGWVRKAARRGMDYYDSVRSVENVIGTSMKD-VL 1268

Qy        847 ITSDQIDKILQDKKDGTV 864
               |   |  : ::     | ||
Db       1269 IGDAQANTLMGQGGDDTV 1286


RESULT 17
FRPC_NEIMB
ID   FRPC_NEIMB     STANDARD;      PRT;  1829 AA.
AC   Q9JYV5;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Iron-regulated protein frpC.
GN   FRPC OR NMB1415.
OS   Neisseria meningitidis (serogroup B).
OC   Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC   Neisseriaceae; Neisseria.
OX   NCBI_TaxID=491;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MC58 / Serogroup B;
RX   MEDLINE=20175755; PubMed=10710307;
RA   Tettelin H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E.,
RA   Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F., Dodson R.J.,
RA   Nelson W.C., Gwinn M.L., DeBoy R., Peterson J.D., Hickey E.K.,
RA   Haft D.H., Salzberg S.L., White O., Fleischmann R.D., Dougherty B.A.,
RA   Mason T., Ciecko A., Parksey D.S., Blair E., Cittone H., Clark E.B.,
RA   Cotton M.D., Utterback T.R., Khouri H., Qin H., Vamathevan J.,
RA   Gill J., Scarlato V., Masignani V., Pizza M., Grandi G., Sun L.,
RA   Smith H.O., Fraser C.M., Moxon E.R., Rappuoli R., Venter J.C.;
RT   "Complete genome sequence of Neisseria meningitidis serogroup B strain
RT   MC58.";
RL   Science 287:1809-1815(2000).
CC   -!- FUNCTION: May participate in the pathogenesis of meningococcal
CC       disease.
CC   -!- SUBCELLULAR LOCATION: Outer-membrane associated and secreted (By
CC       similarity).
```

```
CC   -!- DOMAIN: The Gly-rich region is probably involved in binding
CC       calcium, which is required for target cell-binding or cytolytic
CC       activity.
CC   -!- SIMILARITY: BELONGS TO THE RTX PROKARYOTIC TOXIN FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; AE002490; AAF41776.1; -.
DR   PIR; E81086; E81086.
DR   TIGR; NMB1415; -.
DR   InterPro; IPR001343; Hemlysn_Ca_bind.
DR   InterPro; IPR003995; RtxA.
DR   Pfam; PF00353; hemolysinCabind; 22.
DR   PRINTS; PR00313; CABNDNGRPT.
DR   PRINTS; PR01488; RTXTOXINA.
DR   PROSITE; PS00330; HEMOLYSIN_CALCIUM; 17.
KW   Toxin; Calcium; Outer membrane; Repeat; Complete proteome.
FT   DOMAIN      879   1702       43 X REPEATS, GLY-RICH.
FT   REPEAT      879    884       1.
FT   REPEAT      888    893       2.
FT   REPEAT      897    902       3.
FT   REPEAT     1016   1021       4.
FT   REPEAT     1025   1030       5.
FT   REPEAT     1034   1039       6.
FT   REPEAT     1043   1048       7.
FT   REPEAT     1052   1057       8.
FT   REPEAT     1061   1066       9.
FT   REPEAT     1070   1075       10.
FT   REPEAT     1079   1084       11.
FT   REPEAT     1088   1093       12.
FT   REPEAT     1097   1102       13.
FT   REPEAT     1216   1221       14.
FT   REPEAT     1225   1230       15.
FT   REPEAT     1234   1239       16.
FT   REPEAT     1243   1248       17.
FT   REPEAT     1252   1257       18.
FT   REPEAT     1261   1266       19.
FT   REPEAT     1270   1275       20.
FT   REPEAT     1279   1284       21.
FT   REPEAT     1288   1293       22.
FT   REPEAT     1297   1302       23.
FT   REPEAT     1416   1421       24.
FT   REPEAT     1425   1430       25.
FT   REPEAT     1434   1439       26.
FT   REPEAT     1443   1448       27.
FT   REPEAT     1452   1457       28.
FT   REPEAT     1461   1466       29.
FT   REPEAT     1470   1475       30.
FT   REPEAT     1479   1484       31.
FT   REPEAT     1488   1493       32.
FT   REPEAT     1497   1502       33.
FT   REPEAT     1616   1621       34.
FT   REPEAT     1625   1630       35.
FT   REPEAT     1634   1639       36.
FT   REPEAT     1643   1648       37.
FT   REPEAT     1652   1657       38.
FT   REPEAT     1661   1666       39.
FT   REPEAT     1670   1675       40.
FT   REPEAT     1679   1684       41.
FT   REPEAT     1688   1693       42.
FT   REPEAT     1697   1702       43.
SQ   SEQUENCE   1829 AA;  197445 MW;  8F63506E1F6D9B40 CRC64;

  Query Match             9.4%;  Score 437;  DB 1;  Length 1829;
  Best Local Similarity   20.4%;  Pred. No. 2.6e-14;
  Matches  227;  Conservative  175;  Mismatches  365;  Indels  346;  Gaps   46;

Qy          6 VIKSNIQAGLNSTKSGLKNLYLAIPKDYDPQKGGTLNDFIKAADELGIARLAEEPNHTET 65
              :: :||: |   : |: : :|                :: |   |   | |
Db        243 IVNNNIKQGNEAFKNEINSL--------------VHDMKAAGKEFG------------ 274

Qy         66 AKKSVDTVN-QFLSLTQTGIAISATKLEKFLQKHSTNKLAKGLDSVENIDRKLGKASNVL 124
                   | :| |: :||      :|   :     :  :::  : ||: :::  :   |:  |:: |
Db        275 -----DDLNTQWNNLTQ-----AAEIIYNDIVDNTSQGIEKGVKAIKELSEKMKNAASDL 324

Qy        125 STLSSFLGTALAGIELDSLIKKGDAAPDALAKASIDLINEIIGNLSQSTQTIEAFSSQLA 184
              :               |   :|     : ||:|: :               :: |    |    |
Db        325 A--------------DGSAEKAKQVVEDLAQAAKEAYENAKSTAEKAAQAAREFFKGLP 369

Qy        185 K-------------------------LGSTISQAK------------------GFSNIG 200
                                        |  : : |                    |    :
Db        370 SFKDLAEKFRDLFPNPEGWIDDGHQCLAPWVKETKKRNGKYHVYDPLALDLDGDGIETVA 429

Qy        201 NK-LQNLNFSKTNLGLEIITGLLSGISAGFALADKNASTGKKVAAGFELSNQVIGNVTKA 259
                |      |  || |:    || :|     |   : |  |              | :|  |   :
Db        430 TKGFAGSLFDHTNNGIRTATGWVSA-DDGLLVRDLN-------------GNGIIDNGAEL 475

Qy        260 ISSYVLAQRVAAGLSTTGAVAALITSSIMLAISPLAFMNAADKFNHANALDEFAKQFRKF 319
                  ::| |             |||                       |   :   | |:|   :  : :: :
Db        476 FGDNT---KLADGSFAKHGYAALAELD----------SNGDNIINAADAAFQTLRVWQDL 522

Qy        320 GYDG----DHLLAEYQRGVGTIEASLTTISTALGAVSAGVSAAAVGSAVGTPIALLVAGV 375
                ||    :  |    : |: ::: :     ::   ||      | :  |  ||    |         |
Db        523 NQDGISQANELRTLEELGIQSLDLAYKDVNKNLG---NGNTLAQQGSYTKTD------GT 573

Qy        376 TGLISGILEASKQAMFESVANRLQGKILEWEKQNGGQNYFDKG--YDSRYAAYLANNLKF 433
              |  :  :| |:      ::: :| : |:      :|     |       |      |  | || |: :|
Db        574 TAKMGDLLLAA-----DNLHSRFKDKVELTAEQAKAANLAGIGRLRDLREAAALSGDLAN 628

Qy        434 LSELNKELEA-ERVIAITQ---QRW---DNNIGELA--------------GITKLGERI- 471
              : :      |   |   :|:      :|    |:| |: :                  ||     ::
Db        629 MLKAYSAAETKEAQLALLDNLIHKWAETDSNWGKKSPMRLSTDWTQTANEGIALTPSQVA 688

Qy        472 -----------KSGKAYADAFEDGKKVEAGSNITLDAKTG----------------IIDI 504
                           ||  ||   |         ||| ||                      |: :
Db        689 QLKKNALVSLSDKAKAAIDAARDRIAV-------LDAYTGQDSNTLYYMSEEDALNIVKV 741

Qy        505 SNSN-----------------------------GKKTQALHFTSPLLTAGTESRERLTNG 535
              :|                                   |  |  |  |  |: |       :|  ||
Db        742 TNDTYDHLAKNIYQNLLFQTRLQPYLNQISFKMENDTFTLDF-SGLVQAFNHVKE--TNP 798

Qy        536 KYSYIN---KLKFGRVKNW---------QVTDGEASSKL-DFSKVIQR-----VAETEGT 577
              : ::::   |  :| :::|          |  : :  :  |   |: ||: :        :|:| ||
Db        799 QKAFVDLAEMLAYGELRSWYEGRRLMTDYVEEAKKAGKFEDYQKVLGQETVALLAKTSGT 858

Qy        578 DEIGLIVN------------AKAGNDDIFVGQGKMNIDGGDGHDRVFYSKDGGFGNITV- 624
                  :: |               |||  :   |   |    ::||  |   |      :  :   ||| ||
Db        859 QADDILQNVGFGHNKNVSLYGNDGNDTLIGGAGNDYLEGGSGSDTYVFGE--GFGQDTVY 916

Qy        625 ---------------DGTSA-----TEAGSY-TVNRKVARGDIYHEVVKRQETKVGKRTE 663
                             || :|         |    |::   :    |    :    :    :    |:
Db        917 NYDYATGRKDIIRFTDGITADMLTFTREGNHLLIKAKDGSGQVTVQSYFQNDGSGAYRID 976

Qy        664 TIQYRDYELRKVG----------------YGYQSTDNLK---SVEEVIGSQFNDVFKGSK 704
               |  :  :  ::  |                   |   |||:   |         :  :  |:  :|:  |
Db        977 EIHFDNGKVLDVATVKELVQQSTDGSDRLYAYQSGNTLNGGLGDDYLYGADGDDLLNGDA 1036

Qy        705 FNDIFHSGEGDDLLDGGAGDDRLF---------GGKGNDRLSGDEGDDLLDGGSGDDVLN 755
                ||  :||  |:|  ||||  |:|  |:                ||:|||  |:|::|:|  |  ||:|:|  |
Db       1037 GNDSIYSGNGNDTLDGGEGNDALYGYNGNDALNGGEGNDHLNGEDGNDTLIGGAGNDYLE 1096

Qy        756 GGAGNDVYIFRKGDGNDTLYD---GTG-NDKLAFADANISDIMIERTKEG--IIVKRNDH 809
              ||:|:|  |:|  ||  | ||:|:       ||     |    :   |     :|  |:   |:||   :::|  |
Db       1097 GGSGSDTYVFGKGFGQDTVYNYDYATGRKDIIRFTDGITAD-MLTFTREGNHLLIKAKDG 1155

Qy        810 SGSINIPRWYITSNLQNYQSNKTDHKIEQLIGKDGSYITSDQIDKILQDKKDGTVITSQE 869
              || :  :   ::           ||    |       ::|:::    :|    :        : ::|      ||
Db       1156 SGQVTVQYYF-----QNDGSGA--YRIDEIHFDNGKVLDVATVKELVQQSTDG------- 1201

Qy        870 LKKLADENKSQKLSASDIASSLNKLVGSMALFG 902
```

```
                       | :| |    ::||  :|   |:|
Db       1202 ---------SDRLYAYQSGNTLNGGLGDDYLYG 1225


RESULT 18
FRPC_NEIMC
ID   FRPC_NEIMC     STANDARD;      PRT;  1829 AA.
AC   P55127;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Iron-regulated protein frpC.
GN   FRPC.
OS   Neisseria meningitidis (serogroup C).
OC   Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC   Neisseriaceae; Neisseria.
OX   NCBI_TaxID=135720;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=FAM20 / Serogroup C;
RX   MEDLINE=94018616; PubMed=8412674;
RA   Thompson S.A., Wang L.L., Sparling P.F.;
RT   "Cloning and nucleotide sequence of frpC, a second gene from
RT   Neisseria meningitidis encoding a protein similar to RTX
RT   cytotoxins.";
RL   Mol. Microbiol. 9:85-96(1993).
CC   -!- FUNCTION: May participate in the pathogenesis of meningococcal
CC       disease.
CC   -!- SUBCELLULAR LOCATION: Outer-membrane associated and secreted.
CC   -!- DOMAIN: The Gly-rich region is probably involved in binding
CC       calcium, which is required for target cell-binding or cytolytic
CC       activity.
CC   -!- SIMILARITY: BELONGS TO THE RTX PROKARYOTIC TOXIN FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; L06299; AAA99902.1; -.
DR   PIR; S35027; S35027.
DR   InterPro; IPR001343; Hemlysn_Ca_bind.
DR   InterPro; IPR003995; RtxA.
DR   Pfam; PF00353; hemolysinCabind; 22.
DR   PRINTS; PR00313; CABNDNGRPT.
DR   PRINTS; PR01488; RTXTOXINA.
DR   PROSITE; PS00330; HEMOLYSIN_CALCIUM; 17.
KW   Toxin; Calcium; Outer membrane; Repeat.
FT   DOMAIN      879   1702       43 X REPEATS, GLY-RICH.
FT   REPEAT      879    884       1.
FT   REPEAT      888    893       2.
FT   REPEAT      897    902       3.
FT   REPEAT     1016   1021       4.
FT   REPEAT     1025   1030       5.
FT   REPEAT     1034   1039       6.
FT   REPEAT     1043   1048       7.
FT   REPEAT     1052   1057       8.
FT   REPEAT     1061   1066       9.
FT   REPEAT     1070   1075       10.
FT   REPEAT     1079   1084       11.
FT   REPEAT     1088   1093       12.
FT   REPEAT     1097   1102       13.
FT   REPEAT     1216   1221       14.
FT   REPEAT     1225   1230       15.
FT   REPEAT     1234   1239       16.
FT   REPEAT     1243   1248       17.
FT   REPEAT     1252   1257       18.
FT   REPEAT     1261   1266       19.
FT   REPEAT     1270   1275       20.
FT   REPEAT     1279   1284       21.
FT   REPEAT     1288   1293       22.
FT   REPEAT     1297   1302       23.
FT   REPEAT     1416   1421       24.
FT   REPEAT     1425   1430       25.
FT   REPEAT     1434   1439       26.
FT   REPEAT     1443   1448       27.
FT   REPEAT     1452   1457       28.
FT   REPEAT     1461   1466       29.
FT   REPEAT     1470   1475       30.
FT   REPEAT     1479   1484       31.
FT   REPEAT     1488   1493       32.
FT   REPEAT     1497   1502       33.
FT   REPEAT     1616   1621       34.
FT   REPEAT     1625   1630       35.
FT   REPEAT     1634   1639       36.
FT   REPEAT     1643   1648       37.
FT   REPEAT     1652   1657       38.
FT   REPEAT     1661   1666       39.
FT   REPEAT     1670   1675       40.
FT   REPEAT     1679   1684       41.
FT   REPEAT     1688   1693       42.
FT   REPEAT     1697   1702       43.
SQ   SEQUENCE   1829 AA;  197622 MW;  5C3494020A782DC8 CRC64;

  Query Match              9.4%;  Score 435;  DB 1;  Length 1829;
  Best Local Similarity   21.9%;  Pred. No. 3.2e-14;
  Matches  231;  Conservative  167;  Mismatches  355;  Indels  302;  Gaps   48;

Qy         43 DFIKAAD----ELGIARLAEEPNHTETAKKSVDTVNQFL----SLTQTGIAISATKLEKF 94
              | | ||     : | | |||  :: :    : |   | |     | | ||:
Db        478 DNTKLADGSFAKHGYAALAELDSNGDNIINAADAAFQTLRVWQDLNQDGIS--------- 528

Qy         95 LQKHSTNKLAK-GLDSVE----NIDRKLGKASNVLSTLSSFLGTALAGIELDSLIKKGD- 148
               | :     | : |: |::    :::: ||     | |:    |:  |      :: |:  |
Db        529 -QANELRTLEELGIQSLDLAYKDVNKNLGN-GNTLAQQGSYTKTDGTTAKMGDLLLAADN 586

Qy        149 -----------AAPDALAK--ASIDLINEI---------IGNLSQSTQTIEAFSSQLAKL 186
                         |  |  |   | |  | ::       :  |: ::    |     :|||  |
Db        587 LHSRFKDKVELTAEQAKAANLAGIGRLRDLREAAALSGDLANMLKAYSAAETKEAQLALL 646

Qy        187 GSTISQ-AKGFSNIGNKLQ---NLNFSKT-NLGLEIITGLLSGI--SAGFALADK----- 234
                : |  : |:  || | |      : ::::| | |: :   :: :  :|  :|:||
Db        647 DNLIHKWAETDSNWGKKSPMRLSTDWTQTANEGIALTPSQVAQLKKNALVSLSDKAKAAI 706

Qy        235 ---------NASTGKKVAAGFELSNQVIGNVTKAISSYV--LAQRVAAGLSTTGAVAAL 282
                       :| ||:  :   : :| :    |: |   :      ||: :    |       |
Db        707 DAARDRIAVLDAYTGQDSSTLYYMSEEDALNIVKVTNDTYDHLAKNIYQNL--------L 758

Qy        283 ITSSIMLAISPLAFMNAAD-----------KFNHA------NALDEFAKQFR----KFGY 321
                : :    :: ::|     |            |||        | : |:        :  |
Db        759 FQTRLQPYLNQISFKMENDTFTLDFSGLVQAFNHVKETNPQKAFVDLAEMLAYGELRSWY 818

Qy        322 DGDHLLAEY---QRGVGTIE---ASLTTISTALGAVSAGVSAAAVGSAVGTPIALLVA-- 373
              :|  |:|:|     :    |   |       : || | ::|  |     :   ||      |:
Db        819 EGRRLMADYVEEAKKAGKFEDYQKVLGQETVALLAKTSGTQADDILQNVGFGHNKNVSLY 878

Qy        374 ---GVTGLISGILEASKQAMFESVANRLQGKILEWEKQNGGQNY-FDKG------YDSRY 423
                   |   || |  |   |       :          |  :|   | | ||      |:  |
Db        879 GNDGNDTLIGG---AGNDYL---------------EGGSGSDTYVFGKGFGQDTVYNYDY 920

Qy        424 AAYLANNLKFLSELNKEL-----EAERVI--------AITQQRWDNNIG----------- 459
              |   : ::|   :  ::     |   ::          :| |  :  ||
Db        921 ATGRKDIIRFTDGITADMLTFTREGNHLLIKAKDDSGQVTVQSYFQNDGSGAYRIDEIHF 980

Qy        460 ------ELAGITKLGERIKSGKAYADAFEDGKKVEAG----------SNITLDAKTGIID 503
                    ::| : :| ::  |     |:: |  : |              :  |:   |
Db        981 DNGKVLDVATVKELVQQSTDGSDRLYAYQSGNTLNGGLGDDYLYGADGDDLLNGDAGNDS 1040

Qy        504 ISNSNGKKTQALHFTSPLLTAGTESRERLTNGKYSYINKLKFGRVKNWQVTDGEASSKLD 563
              | : || |          | | :  |                  :|   |  :  |
Db       1041 IYSGNGNDT----------LNGGEGNDAL------------YGYNGNDALNGG------- 1071

Qy        564 FSKVIQRVAETEGTDEIGLIVNAKAGNDDIFVGQGKMNIDGGDGHDRVFYSKDGGFGNIT 623
```

```
Db      1072 -----------EGNDHL----NGEDGNDTLIGGAGNDYLEGGSGSDTYVFGK--GFGQDT 1114

Qy       624 V----------------DGTSA-----TEAGSY-TVNRKVARGDIYHEVVKRQETKVGKR 661

Db      1115 VYNYDYATGRKDIIRFTDGITADMLTFTREGNHLLIKAKDGSGQVTVQSYFQNDGSGAYR 1174

Qy       662 TETIQYRDYELRKVG----------------YGYQSTDNLK---SVEEVIGSQFNDVFKG 702

Db      1175 IDEIHFDNGKVLDVATVKELVQQSTDGSDRLYAYQSGNTLNGGLGDDYLYGADGDDLLNG 1234

Qy       703 SKFNDIFHSGEGDDLLDGGAGDDRLF---------GGKGNDRLSGDEGDDLLDGGSGDDV 753

Db      1235 DAGNDSIYSGNGNDTLDGGEGNDALYGYNGNDALNGGEGNDHLNGEDGNDTLIGGAGNDY 1294

Qy       754 LNGGAGNDVYIFRKGDGNDTLYD---GTG-NDKLAFADANISDIMIERTKEG--IIVKRN 807

Db      1295 LEGGSGSDTYVFGKGFGQDTVYNYDYATGRKDIIRFTDGITAD-MLTFTREGNHLLIKAK 1353

Qy       808 DHSGSINIPRWYITSNLQNYQSNKTDHKIEQLIGKDGSYITSDQIDKILQDKKDGTVITS 867

Db      1354 DDSGQVTVQSYF-----QNDGSGA--YRIDEIHFDNGKVLDVATVKELVQQSTDG----- 1401

Qy       868 QELKKLADENKSQKLSASDIASSLNKLVGSMALFG 902

Db      1402 -----------SDRLYAYQSGSTLNGGLGDDYLYG 1425
```

```
RESULT 19
FRPA_NEIMB
ID   FRPA_NEIMB     STANDARD;      PRT;  1302 AA.
AC   Q9K0K9;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Iron-regulated protein frpA.
GN   FRPA OR NMB0585.
OS   Neisseria meningitidis (serogroup B).
OC   Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC   Neisseriaceae; Neisseria.
OX   NCBI_TaxID=491;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MC58 / Serogroup B;
RX   MEDLINE=20175755; PubMed=10710307;
RA   Tettelin H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E.,
RA   Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F., Dodson R.J.,
RA   Nelson W.C., Gwinn M.L., DeBoy R., Peterson J.D., Hickey E.K.,
RA   Haft D.H., Salzberg S.L., White O., Fleischmann R.D., Dougherty B.A.,
RA   Mason T., Ciecko A., Parksey D.S., Blair E., Cittone H., Clark E.B.,
RA   Cotton M.D., Utterback T.R., Khouri H., Qin H., Vamathevan J.,
RA   Gill J., Scarlato V., Masignani V., Pizza M., Grandi G., Sun L.,
RA   Smith H.O., Fraser C.M., Moxon E.R., Rappuoli R., Venter J.C.;
RT   "Complete genome sequence of Neisseria meningitidis serogroup B strain
RT   MC58.";
RL   Science 287:1809-1815(2000).
CC   -!- FUNCTION: May participate in the pathogenesis of meningococcal
CC       disease.
CC   -!- SUBCELLULAR LOCATION: Outer-membrane associated and secreted (By
CC       similarity).
CC   -!- DOMAIN: The Gly-rich region is probably involved in binding
CC       calcium, which is required for target cell-binding or cytolytic
CC       activity.
CC   -!- SIMILARITY: BELONGS TO THE RTX PROKARYOTIC TOXIN FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; AE002414; AAF41013.1; -.
DR   PIR; C81182; C81182.
DR   TIGR; NMB0585; -.
DR   InterPro; IPR001343; Hemlysn_Ca_bind.
DR   InterPro; IPR003995; RtxA.
DR   Pfam; PF00353; hemolysinCabind; 12.
DR   PRINTS; PR00313; CABNDNGRPT.
DR   PRINTS; PR01488; RTXTOXINA.
DR   PROSITE; PS00330; HEMOLYSIN_CALCIUM; 9.
KW   Toxin; Calcium; Outer membrane; Repeat; Complete proteome.
FT   DOMAIN      752    975       13 X REPEATS, GLY-RICH.
FT   REPEAT      761    766       2.
FT   REPEAT      770    775       3.
FT   REPEAT      889    894       4.
FT   REPEAT      898    903       5.
FT   REPEAT      907    912       6.
FT   REPEAT      916    921       7.
FT   REPEAT      925    930       8.
FT   REPEAT      934    939       9.
FT   REPEAT      943    948       10.
FT   REPEAT      952    957       11.
FT   REPEAT      961    966       12.
FT   REPEAT      970    975       13.
SQ   SEQUENCE   1302 AA;  141397 MW;  21D058C56C98BDE8 CRC64;

  Query Match             9.1%;  Score 421;  DB 1;  Length 1302;
  Best Local Similarity   22.5%;  Pred. No. 1e-13;
  Matches  198;  Conservative  136;  Mismatches  292;  Indels  254;  Gaps   39;

Qy       193 AKGFSNIGNKLQNLNFSKTNLGLEIITGLLSGISAGFALADKNASTGKKVAAGFELSNQV 252

Db       303 AKGFS-------GSLFDHTNNGIRTATGWVSA-DDGLLVRDLN-------------GNGI 341

Qy       253 IGNVTKAISSYVLAQRVAAGLSTTGAVAALITSSIMLAISPLAFMNAADKFNHANALDEF 312

Db       342 IDNGAELFGDNT---KLADGSFAKHGYAALAELD----------SNGDNIINAADAAFQS 388

Qy       313 AKQFRKFGYDG----DHLLAEYQRGVGTIEASLTTISTALGAVSAGVSAAAVGSAVGTPI 368

Db       389 LRVWQDLNQDGISQANELRTLEELGIQSLDLAYKDVNKNLG---NGNTLAQQGSYTKTD- 444

Qy       369 ALLVAGVTGLISGILEASKQAMFESVANRLQGKILEWEKQNGGQNYFDKG--YDSRYAAY 426

Db       445 -----GTTAKMGDLLLAA-----DNLHSRFKDKVELTAEQAKAANLAGIGRLRDLREAAA 494

Qy       427 LANNLKFLSELNKELEA-ERVIAITQ---QRW---DNNIGELA--------------GIT 465

Db       495 LSGDLANMLKAYSAAETKEAQLALLDNLIHKWAETDSNWGKKSPMRLSTDWTQTANEGIA 554

Qy       466 KLGERI------------KSGKAYADAFEDGKKVEAGSNITLDAKTG------------- 500

Db       555 LTPSQVAQLKKNALVSLSDKAKAAIDAARDRIAV-------LDAYTGQDSNTLYYMSEED 607

Qy       501 ---IIDISNSN----------------------------GKKTQALHFTSPLLTAGTES 528

Db       608 ALNIVKVTNDTYDHLAKNIYQNLLFQTRLQPYLNQISFKMENDTFTLDF-SGLVQAFNHV 666

Qy       529 RERLTNGKYSYIN---KLKFGRVKNW---------QVTDGEASSKL-DFSKVIQR----- 570

Db       667 KE--TNPQKAFVDLAEMLAYGELRSWYEGRRLMTDYVEEAKKAGKFEDYQKVLGQETVAL 724

Qy       571 VAETEGTDEIGLIVN------------AKAGNDDIFVGQGKMNIDGGDGHDRVFYSKDGG 618

Db       725 LAKTSGTQADDILQNVGFGHNKNVSLYGNDGNDTLIGGAGNDYLEGGSGSDTYVFGE--G 782

Qy       619 FGNITV----------------DGTSA-----TEAGSY-TVNRKVARGDIYHEVVKRQET 656

Db       783 FGQDTVYNYDYATGRKDIIRFTDGITADMLTFTREGNHLLIKAKDGSGQVTVQSYFQNDG 842

Qy       657 KVGKRTETIQYRDYELRKVG----------------YGYQSTDNLK---SVEEVIGSQFN 697

Db       843 SGAYRIDEIHFDNGKVLDVATVKELVQQSTDGSDRLYAYQSGNTLNGGLGDDYLYGADGD 902
```

```
Qy        698 DVFKGSKFNDIFHSGEGDDLLDGGAGDDRLF---------GGKGNDRLSGDEGDDLLDGG 748
              |:  |   ||  :|| |:| |||| |:| |:          ||:||| |:|::|:| | ||
Db        903 DLLNGDAGNDSIYSGNGNDTLDGGEGNDALYGYNGNDALNGGEGNDHLNGEDGNDTLIGG 962

Qy        749 SGDDVLNGGAGNDVYIFRKGDGNDTLYD---GTG-NDKLAFADANISDIMIERTKEG--I 802
              :|:| | ||:|:| |:| || | | :|:     ||  | : | |    :| |:  |:|| :
Db        963 AGNDYLEGGSGSDTYVFGKGFGQDAVYNYDYATGRKDIIRFTDGITAD-MLTFTREGNHL 1021

Qy        803 IVKRNDHSGSINIPRWYITSNLQNYQSNKTDHKIEQLIGKDGSYITSDQIDKILQDKKDG 862
              ::|  | || : :   ::         || |    ::|:::  :|  :    :::|  ||
Db       1022 LIKAKDGSGQVTVQSYF-----QNDGSGA--YRIDEIHFDNGKVLDVATVKELVQQSTDG 1074

Qy        863 TVITSQELKKLADENKSQKLSASDIASSLNKLVGSMALFG 902
                                    |:| |   ::||  :|   |:|
Db       1075 ----------------SDRLYAYQSGNTLNGGLGDDYLYG 1098
```

RESULT 20
FRPA_NEIMC

```
ID   FRPA_NEIMC     STANDARD;      PRT;  1115 AA.
AC   P55126;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Iron-regulated protein frpA.
GN   FRPA.
OS   Neisseria meningitidis (serogroup C).
OC   Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC   Neisseriaceae; Neisseria.
OX   NCBI_TaxID=135720;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=FAM20 / Serogroup C;
RX   MEDLINE=93139051; PubMed=8423153;
RA   Thompson S.A., Wang L.L., West A., Sparling P.F.;
RT   "Neisseria meningitidis produces iron-regulated proteins related to
RT   the RTX family of exoproteins.";
RL   J. Bacteriol. 175:811-818(1993).
CC   -!- FUNCTION: May participate in the pathogenesis of meningococcal
CC       disease.
CC   -!- SUBCELLULAR LOCATION: Outer-membrane associated and secreted.
CC   -!- DOMAIN: The Gly-rich region is probably involved in binding
CC       calcium, which is required for target cell-binding or cytolytic
CC       activity.
CC   -!- SIMILARITY: BELONGS TO THE RTX PROKARYOTIC TOXIN FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; L06302; AAA25454.1; -.
DR   PIR; A47058; A47058.
DR   InterPro; IPR001343; Hemlysn_Ca_bind.
DR   InterPro; IPR003995; RtxA.
DR   Pfam; PF00353; hemolysinCabind; 7.
DR   PRINTS; PR00313; CABNDNGRPT.
DR   PRINTS; PR01488; RTXTOXINA.
DR   PROSITE; PS00330; HEMOLYSIN_CALCIUM; 5.
KW   Toxin; Calcium; Outer membrane; Repeat.
FT   DOMAIN      765    988       13 X REPEATS, GLY-RICH.
FT   REPEAT      774    779       2.
FT   REPEAT      783    788       3.
FT   REPEAT      902    907       4.
FT   REPEAT      911    916       5.
FT   REPEAT      920    925       6.
FT   REPEAT      929    934       7.
FT   REPEAT      938    943       8.
FT   REPEAT      947    952       9.
FT   REPEAT      956    961       10.
FT   REPEAT      965    970       11.
FT   REPEAT      974    979       12.
FT   REPEAT      983    988       13.
SQ   SEQUENCE   1115 AA;  122177 MW;  21841065BCCACEFA CRC64;

  Query Match             8.8%;  Score 408.5;  DB 1;  Length 1115;
  Best Local Similarity   20.4%;  Pred. No. 3.6e-13;
  Matches  239;  Conservative  171;  Mismatches  364;  Indels  399;  Gaps   51;

Qy         40 TLNDFIKAADELGIARLAEEPNHTETAKKSVDTVNQFLSL-------TQTG--------- 83
              |:|: |:|||   :: :  :    | ||        | ||:|       :||
Db         55 TVNNAIQAADSF-LSSIGRDNKITNTASLLASLDNIFLNLRNVSRDIRETGKFKPNDIQQ 113

Qy         84 ----IAISA--------TKLEKFLQ-KHSTNKLAKGLDSVENIDRKLGKASNVLSTLSSF 130
                  | |:|          : |   |  |    ||  ||: |   :: :: |:| ||    ::
Db        114 AIGDIFIAAGDGLQYIKQQTEAMAQSKFLPTKLKTGLNDV--LNSRMLKSSTVLQHELNY 171

Qy        131 LGTALA--GIEL--DSLIKKGDAAPDALAK--ASIDLINEIIGNLSQSTQTIEAFSSQLA 184
              ||  :    | |    :|::     |   |: |    |   |  : ::  :|:   ::    :
Db        172 LGFKIKDYGNERLGESIMNIDDFTPSKIANFFADPDTYSNVLEEVSRFIYSLVPDDANPW 231

Qy        185 K-----LGSTISQ------------------------------------------------ 192
              |     :|  ||:
Db        232 KGGEDYIGRGISEWGELLEKWYKQDFLPYLEKEWDQFPKFEDWLPEFPEWAREWLKLDPK 291

Qy        193 ------------------------AKGFSNIGNKLQNLNFSKTNLGLEIITGLLSGISAG 228
                                      ||||:        |    | |:   || :|    |
Db        292 RSGKYHVYDPLALDLDGDGIETVAAKGFA-------GALFDHRNQGIRTATGWVSA-DDG 343

Qy        229 FALADKNASTGKKVAAGFELSNQVIGNVTKAISSYVLAQRVAAGLSTTGAVAALITSSIM 288
                : | |             | :| |  :         ::| |        |||
Db        344 LLVRDLN-------------GNGIIDNGAELFGDNT---KLADGSFAKHGYAALAELD-- 385

Qy        289 LAISPLAFMNAADKFNHANALDEFAKQFRKFGYDG----DHLLAEYQRGVGTIEASLTTI 344
                        |  :  | |:|  :  : ::       ||      : |    : |: ::: :    :
Db        386 --------SNGDNIINAADAAFQTLRVWQDLNQDGISQANELRTLEELGIQSLDLAYKDV 437

Qy        345 STALGAVSAGVSAAAVGSAVGTPIALLVAGVTGLISGILEASKQAMFESVANRLQGKILE 404
              :  ||    |  :  |  ||          |  |  :  | |:      ::: :| |:
Db        438 NKNLG---NGNTLAQQGSYTKTD------GTTAKMGDLLLAA-----DNLHSRFKDKVEL 483

Qy        405 WEKQNGGQNYFDKG--YDSRYAAYLANNLKFLSELNKELEA-ERVIAITQ---QRW---D 455
                :|   |    |    |  |  | || |: :|  : :         |   |  :|:     :|  |
Db        484 TAEQAKAANLAGIGRLRDLREAAALSGDLANMLKAYSAAETKEAQLALLDNLIHKWAETD 543

Qy        456 NNIGELA--------------GITKLGERI------------KSGKAYADAFEDGKKVEA 489
              :| |: :                 ||     ::                  ||  ||  |   |
Db        544 SNWGKKSPMRLSTDWTQTANEGIALTPSQVAQLKKNALVSLSDKAKAAIDAARDRIAV-- 601

Qy        490 GSNITLDAKTG----------------IIDISNSN------------------------- 508
                   ||| ||                |: ::|
Db        602 -----LDAYTGQDSSTLYYMSEEDALNIVKVTNDTYDHLAKNIYQNLLFQTRLQPYLNQI 656

Qy        509 ----GKKTQALHFTSPLLTAGTESRERLTNGKYSYIN---KLKFGRVKNW---------Q 552
                   |  |  |  |  | |: |       :|    ||  : ::::     | :| :::|
Db        657 SFKMENDTFTLDF-SGLVQAFNHVKE--TNPQKAFVDLAEMLAYGELRSWYEGRRLMADY 713

Qy        553 VTDGEASSKL-DFSKVIQR-----VAETEGTDEIGLIVN------------AKAGNDDIF 594
              | : : : |  |: ||: :       :|:| ||     :: |                ||| :
Db        714 VEEAKKAGKFEDYQKVLGQETVALLAKTSGTQADDILQNVGFGHNKNVSLYGNDGNDTLI 773

Qy        595 VGQGKMNIDGGDGHDRVFYSKDGGFGNITV-----------------DGTSA-----TEAG 633
               | |     ::|| | |     : |   ||| ||                || :|      |  |
Db        774 GGAGNDYLEGGSGSDTYVFGK--GFGQDTVYNYDYATGRKDIIRFTDGITADMLTFTREG 831

Qy        634 SY-TVNRKVARGDIYHEVVKRQETKVGKRTETIQYRDYELRKVG---------------- 676
              :: :   |     |  :    :  | :  | : : :: |
Db        832 NHLLIKAKDDSGQVTVQSYFQNDGSGAYRIDEIHFDNGKVLDVATVKELVQQSTDGSDRL 891

Qy        677 YGYQSTDNLK---SVEEVIGSQFNDVFKGSKFNDIFHSGEGDDLLDGGAGDDRLF----- 728
              | |||     |        : : |:  ||:  |    :||  ||| |||| |:| |:
Db        892 YAYQSGSTLNGGLGDDYLYGADGNDLLNGDAGNDSIYSGNGNDTLDGGEGNDALYGYNGN 951
```

```
Qy        729 ----GGKGNDRLSGDEGDDLLDGGSGDDVLNGGAGNDVYIFRKGDGNDTLYD---GTGND 781
                  ||:||| |:|::|:| | ||:|:| | ||:|:| |:| :| | ||:|:       :|
Db        952 DALNGGEGNDHLNGEDGNDTLIGGAGNDYLEGGSGSDTYVFGEGFGQDTVYNYHVDKNSD 1011

Qy        782 KLAFADANISDIMIERTKEGIIVKRNDHSGSINIPRWYITSNLQNYQSNKTDHKIEQLIG 841
               : |     :|:   |:   :::  ::   :: |  ::  |            |:::  :
Db       1012 TMHFKGFKAADVHFIRSGSDLVLSASEQD-NVRISGFFYGEN----------HRVDTFVF 1060

Qy        842 KDGSYITSDQIDKILQDKKDGTVITSQELKKLADENKSQKLSASDIASSL---NKLVGSM 898
               | :                                      :|  | |  :    | || ||
Db       1061 DDAA-----------------------------------ISNPDFAKYINAGNNLVQSM 1084

Qy        899 ALFG-----TANSVSSNALQPITQPTQGILAPS 926
              ::||      |  :| :| :| : ||   :: ||
Db       1085 SVFGSNTAATGGNVDAN-IQSVQQPL--LVTPS 1114
```

RESULT 21
HLYA_SERMA
ID HLYA_SERMA STANDARD; PRT; 1608 AA.
AC P15320;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 01-NOV-1990 (Rel. 16, Last annotation update)
DE Hemolysin precursor.
GN SHLA.
OS Serratia marcescens.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Serratia.
OX NCBI_TaxID=615;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 31-40.
RC STRAIN=SN8;
RX MEDLINE=88257037; PubMed=3290200;
RA Poole K., Schiebel E., Braun V.;
RT "Molecular characterization of the hemolysin determinant of Serratia
RT marcescens.";
RL J. Bacteriol. 170:3177-3188(1988).
CC -!- FUNCTION: BACTERIAL HEMOLYSINS ARE EXOTOXINS THAT ATTACK BLOOD
CC CELL MEMBRANES AND CAUSE CELL RUPTURE BY MECHANISMS NOT CLEARLY
CC DEFINED.
CC -!- FUNCTION: CELL-BOUND HEMOLYSIN, WHICH RELEASES HEME-IRON FROM
CC ERYTHROCYTES BY INTERACTION WITH THE ERYTHROCYTE MEMBRANE. SHLA
CC REQUIRES SHLB FUNCTION.
CC -!- SUBCELLULAR LOCATION: Outer membrane.
CC -!- SIMILARITY: TO P.MIRABILIS HEMOLYSIN (HPMA).
CC ---

CC ---
DR EMBL; M22618; AAA50323.1; -.
DR PIR; A28182; A28182.
KW Hemolysis; Toxin; Outer membrane; Signal.
FT SIGNAL 1 30
FT CHAIN 31 1608 HEMOLYSIN.
SQ SEQUENCE 1608 AA; 165078 MW; D669B476FE7DAD51 CRC64;

Query Match 4.1%; Score 189.5; DB 1; Length 1608;
Best Local Similarity 20.9%; Pred. No. 0.058;
Matches 220; Conservative 123; Mismatches 406; Indels 305; Gaps 48;

```
Qy         11 IQAGLNSTKSGLK---NLYLAIPKDYDPQKGGTLNDFIKAADE-------LGIARLAEEP 60
              ::|| | :  | |         | |  |  | |      || :           :|
Db        493 LKAGRNVSTQGAKVHAQRDLTIDADNQIQVGVQKTANAKAVRDDKTSWGGIGGGDNKNNS 552

Qy         61 NHTETAKKSVDTVNQFLSLT-QTGIAISATKLE-----KFLQKHSTNKLAKGLD-SVENI 113
              | | :  |  |     | |  | |: |: :|        :     |   ::     | :|: |
Db        553 NRREISHASELTSGGTLRLNGQQGVTITGSKARGQKGGEVTATHGGLRIDNALSTTVDKI 612

Qy        114 DRKLGKASNVLSTL----SSFLGTALAGIELDSLI-----KKGDAAPDALAKASIDLINE 164
               | : | | |: |: |:       :|:  :  : :: |: :    |   |     :|      :
Db        613 DARTGTAFNITSSSHKADNSYQSSTASELKSDTNLTLVSHKDADVIGSQVASGGELSVES 672

Qy        165 IIGNLS------------QSTQTIEAFSSQLA----------------KLGSTISQAK 194
               ||::                 :: |:  :: :                    |  | : |
Db        673 KTGNINVKAAERQQNIDEQKTALTVNGYAKEAGDKQYRAGLRIEHTRDSEKTTRTENSAS 732

Qy        195 GFSNIGNKL---QNLNFSKTNLGLEIITGLLSGISAGFALA-DKNAST----------- 238
               |   ||   ::: || : |  :      :||    |  | || ||
Db        733 SLSGGSVKLKAEKDVTFSGSKLVADKGDASVSGNKVSFLAADDKTASNTEQTKIGGGFYY 792

Qy        239 -------GKKVAAGFELSNQVIGNVTKAISSYVLAQRVAAGLSTTGAVAALITSSIMLAI 291
                      |  | ||:|        |  |:| ||      : : :|     | :
Db        793 TGGIDKLGSGVEAGYE------NNKTQAQSS----KAITSGSDVKGNLT---------- 831

Qy        292 SPLAFMNAADKFNHANALDEFAKQFRKFGYDGDHLLAEYQRGVGTIEASLTTISTALGA- 350
                   :|| ||       |         :::          ||| |        || || | :|
Db        832 -----INARDKLTQQGAQHSVGGAYQENAAGVDHLAA-------ADTASTTTTKTDVGVN 879

Qy        351 VSAGVSAAA----VGSAVGTPIALLVAGVTGLISGI-----------------LEASKQ 388
              : |  |   :|       | |||       ||     | ||                   :| |
Db        880 IGANVDYSAVTRPVERAVGKAAKLDATGVINDIGGIGAPNVGLDIGAQGGSSEKRSSSQ 939

Qy        389 AMFESVAN-----RLQGKILEWEKQNGGQNYFDKG-----YDSRYAAYLANNLKFLSELN 438
              |: ||            :|::     :  |   |        ||     || :    ||     |
Db        940 AVVSSVQAGSIDINAKGEV----RDQGTQYQASKGAVNLTADSHRSEAAANRQDEQSRDT 995

Qy        439 KELEAERVIAIT-------------QQRWDNN-----IGELAGITKLGERIKSGKAYADA 480
               :    ||    |                   || :::        | :     :    :|    |
Db        996 RGSAGVRVYTTTGSDLTVDAKGEGGTQRSNSSASQAVTGSIDAANGINVNVKKDAIYQGT 1055

Qy        481 FEDGKK----VEAGSNITLDAKTGIIDISNS--NGKKTQALHFTSPLLTAGTESRERLTN 534
                :| :      | || :| ||  :       |   |  |  :     ||:        |       |
Db       1056 ALNGGRGKTAVNAGGDIRLDQASDKQSESRSGFNVKASAKGGFTADSKNFGAGFGGGTHN 1115

Qy        535 GKYSYIN------------KLKFGRVKNWQVTDGEASSKLDFS---KVIQRVAETEGTDE 579
              |: |                 :|| ||      | ||  :: |   :  | ||   ||: |  :
Db       1116 GESSSSTAQVGNISGQQGVELKAGRDLTLQGTDVKSQGDVSLSAGNKVALQAAESTQTRK 1175

Qy        580 IGLIVNAKAGNDDIFVGQGKMNIDGGDGHDRVFYSKDGGFGNITVDGT--------SATE 631
                :    :|| |: |                     ||:   ||::   |          ||||
Db       1176 ESKL----SGNIDLGAGSSD-------------SKEKTGGNLSAGGAFDIAKVNESATE 1217

Qy        632 AGSYTV--NRKV-----ARGDIYHEVVKRQETKVGKRTETIQYRD---------YELRKV 675
                   |:  : ||          :||    : :   |   ||     :  :: ::              | |
Db       1218 RQGATIASDGKVTLSANGKGD---DALHLQGAKVSGGSAALEAKNGGILLESAKNEQHKD 1274

Qy        676 GYGYQSTDNLKSVEEVIGSQFNDVFKGSKFNDIFHSGEGDDLLDGG------------- 721
                :   |         |   ||     |  | :   ::|:    |   |
Db       1275 NWSLGIKANAKG-----GQTFNK-DAGGKVDP--NTGKDTHTLGAGLKVGVEQQDKTTHA 1326

Qy        722 -----AGDDRLFGGKGNDRLSGDEGDDLLDGGSGDDVLNGGAGNDVYIFRKGDGNDTLYD 776
                    ||| |   || ||  :   ||:|     |    :  |  |:::  :  :    : :
Db       1327 NTGITAGDVTLNSGK-DTRLAGARVD--------ADSVQGKVGGDLHVESRKD----VEN 1373

Qy        777 GTGNDKLAFADANISDIMIERTKEGIIVKRNDHSGSINIPRWYITSNLQNY----QSNKT 832
              |    |    || :|                  ||   |  |  |||  |             : |
Db       1374 GVKVD----VDAGLS-----------------HS---NDPGSSITSKLSKVGTPRYAGKV 1409

Qy        833 DHKIEQLIGKDGSYITSDQIDKI---LQDKKDGT 863
                |:|  : |     |:|: : :    |   ::| |
Db       1410 KEKLEAGVNKVAD-ATTDKYNSVARRLDPQQDTT 1442
```

RESULT 22
RBP1_PLAVB
ID RBP1_PLAVB STANDARD; PRT; 2869 AA.
AC Q00798;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)

```
DE   Reticulocyte binding protein 1 precursor.
GN   RBP1.
OS   Plasmodium vivax (strain Belem).
OC   Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX   NCBI_TaxID=31273;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92315338; PubMed=1617731;
RA   Galinski M.R., Medina C.C., Ingravallo P., Barnwell J.W.;
RT   "A reticulocyte-binding protein complex of Plasmodium vivax
RT   merozoites.";
RL   Cell 69:1213-1226(1992).
CC   -!- FUNCTION: INVOLVED IN RETICULOCYTE ADHESION. SPECIFICALLY BINDS TO
CC       HUMAN RETICULOCYTE CELLS.
CC   -!- SUBUNIT: Homodimer (Potential).
CC   -!- SUBCELLULAR LOCATION: Membrane-bound.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; M88097; AAA29743.1; -.
KW   Malaria; Receptor; Signal; Transmembrane.
FT   SIGNAL        1     17       POTENTIAL.
FT   CHAIN        18   2869       RETICULOCYTE BINDING PROTEIN 1.
FT   DOMAIN       18   2807       EXTRACELLULAR.
FT   TRANSMEM   2808   2826       POTENTIAL.
FT   DOMAIN     2827   2869       CYTOPLASMIC.
FT   SITE       1030   1032       CELL ATTACHMENT SITE (POTENTIAL).
FT   SITE       2599   2601       CELL ATTACHMENT SITE (POTENTIAL).
SQ   SEQUENCE   2869 AA;  330213 MW;  B9DBE442205EBCFF CRC64;

  Query Match              4.1%;  Score 188.5;  DB 1;  Length 2869;
  Best Local Similarity   19.7%;  Pred. No. 0.14;
  Matches  209;  Conservative  175;  Mismatches  360;  Indels  315;  Gaps   50;

Qy          2 SNINVIKSNIQA--------GLNSTKSGLKNLYLAIPKDYDPQKGGTLNDFIKAADELGI 53
Db       1214 STANTLKSNAKGKENEHDLEELNKTKGQMRDIYEKLKKIAEELKEGTVNE-LKDANEKA- 1271

Qy         54 ARLAEEPNHT-------------ETAKKSVDTVNQFLSLTQTGIAISATKLEKFLQKHS- 99
Db       1272 NKVEPEPERNIIGHVLERITVEKDKAGKVVEEMNSL-----------KTKIEKLIQETSD 1320

Qy        100 -------TNKLAKGLDSV---ENIDRKLGKASNVLSTLSSFLGTALAGIELDSLIKKGDA 149
Db       1321 DSQNELVTTSITKHLENAKGYEDVIKRNEEDSIQLREKAKSLET------LDEMKK---- 1370

Qy        150 APDALAKASIDLINEIIGNLSQSTQTIEAFSSQLAKLGSTISQ--AKGFSNI-----GNK 202
Db       1371 -----------LVQQVNMNLQSAIQGNAGISKELNELKGVIELLISTNYSSILEYVKKNS 1419

Qy        203 LQNLNFSKTNLGLEIITGLLSGISAGFALADKNASTGKKVAAGFELSNQVIGNVTKAISS 262
Db       1420 SESVRFSQ----------LANGEFTKAEGEEKNASA--RLAEAEKLKEQIV----KDLDY 1463

Qy        263 YVLAQRVAAGLSTTGAVAALITSSIMLAISPLAFMNAADKF-----NHANALDEFAK--Q 315
Db       1464 SDIDDKVK-------KIEGIKREILKMKESALTFWEESEKFKQMCSSHMENAKEGKKKIE 1516

Qy        316 FRKFGYDG------DHLLAEYQRGVGTIEASLTTISTALGAVSAGVSAAAVGSAVGTPIA 369
Db       1517 YLKNNGDGGKANITDSQMEEVGNYVSKAEHAFHTVEAQVDKTKAFCES------------ 1564

Qy        370 LLVAGVTGLISGILEASKQAMFESVANRLQGKILEWEKQNGGQNYFD---KGYDSRYAAY 426
Db       1565 -IVAYVTKM--------DNLFNESLMKEVKVKC---EKKNDEAEKYSAKLKPYDGRIKAR 1612

Qy        427 LANNLKFLSELNKELEAER---------------VIAITQQRWD---NNIG--------- 459
Db       1613 VSENERKISELKEKAKVEKKESSQLNDVSTKSLLQIDNCRQQLDSVLSNIGRVKQNALQY 1672

Qy        460 ---------ELAGITKLG-----ERIKSGK-AYADAFEDGK------KVEAGSNITLDAK 498
Db       1673 FDSADKSMKSVLPISELGAEKSLDKVKAAKESYEKNLETVQNEMSRINVEEGSLTDIDKK 1732

Qy        499 TGIIDISNSNGKKTQALHFTSPLLTAGTESRERLTNGKYSYINKLKFG---RVKNWQVTD 555
Db       1733 --ITDIEND-------------LLKMKKQYEEGL-------LQKIKENADKRKSNFELVG 1770

Qy        556 GEASSKLDFSKVIQRVAETEGTDEIGLIVNAKAGNDDIFVGQGKMNIDGGDGHDRVFYSK 615
Db       1771 SEINALLDPSTSIFIKLKLKEYDMTGDLKNYGV----------KMNEIHGE------FTK 1814

Qy        616 DGGFGNITVDGTSATEAGSYTVNRKVARGDIYHEVVKRQETKVGKRTETIQYRDYELRKV 675
Db       1815 -----SYNLIETHLSNATDYSVTFEKAQS--LRELAEKEEEHLRRREEEAIFLLNDIKKV 1867

Qy        676 GYGYQSTDNLKSVEEVIGSQFNDVFKGSKFNDIFHSGEGDDLLDGGAGDDRLFGGKGNDR 735
Db       1868 -------ESLKLLKEMM-KKVSAEYEGMKRDHTSVSQLVQDM-----------------K 1902

Qy        736 LSGDEGDDLLDGGSGDDVLNGGAGNDVYIFRK----------GDGNDTLYDGTGNDKLAF 785
Db       1903 TIVDELKTLNDISECSSVLN----NVVSIVKKVKESKHADYRRDAN-SMYE----SMVTL 1953

Qy        786 ADANISD-IMIERTKEGIIVKRNDHSGSINIPRWYITSNLQNYQSNKTDHKIEQ------ 838
Db       1954 ANYFLSDEAKISSGMEFNAEMKSNFKTDLELEIFSVISN-----SNELLKKIEQDSNDVI 2008

Qy        839 -------LIGKDGSYI-----TSDQIDKILQDKKDGTVITSQE----LKKLAD------- 875
Db       2009 QKERESEQLAKDATDIYNVIKLKNEFNEKLEEAKNKEEVVSEKVREALKRLSQVEGIRCH 2068

Qy        876 -ENKSQKLSASDIASSLNKLVGSMALFGTANSVSSNALQ 913
Db       2069 FENFHRLLDNTEELENLKKMV---TIYRDKKSERESGLQ 2104

RESULT 23
HLYA_PROMI
ID   HLYA_PROMI     STANDARD;      PRT;  1577 AA.
AC   P16466;
DT   01-AUG-1990 (Rel. 15, Created)
DT   01-AUG-1990 (Rel. 15, Last sequence update)
DT   01-NOV-1990 (Rel. 16, Last annotation update)
DE   Hemolysin precursor.
GN   HPMA.
OS   Proteus mirabilis.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Proteus.
OX   NCBI_TaxID=584;
RN   [1]
RP   SEQUENCE FROM N.A., AND SEQUENCE OF 30-43.
RC   STRAIN=Isolate 477-12;
RX   MEDLINE=90170827; PubMed=2407716;
RA   Uphoff T.S., Welch R.A.;
RT   "Nucleotide sequencing of the Proteus mirabilis calcium-independent
RT   hemolysin genes (hpmA and hpmB) reveals sequence similarity with the
RT   Serratia marcescens hemolysin genes (shlA and shlB).";
RL   J. Bacteriol. 172:1206-1216(1990).
CC   -!- FUNCTION: BACTERIAL HEMOLYSINS ARE EXOTOXINS THAT ATTACK BLOOD
CC       CELL MEMBRANES AND CAUSE CELL RUPTURE BY MECHANISMS NOT CLEARLY
CC       DEFINED.
CC   -!- FUNCTION: CELL-BOUND HEMOLYSIN, WHICH RELEASES HEME-IRON FROM
CC       ERYTHROCYTES BY INTERACTION WITH THE ERYTHROCYTE MEMBRANE. HPMA
CC       REQUIRES HPMB FUNCTION.
CC   -!- SUBCELLULAR LOCATION: Outer membrane.
CC   -!- MISCELLANEOUS: THE CONSERVED AMPHIPATHIC DOMAINS IN SHLA AND HPMA
CC       MAY BE RESPONSIBLE FOR PORE FORMATION.
CC   -!- SIMILARITY: TO S.MARCESCENS HEMOLYSIN (SHLA).
CC   --------------------------------------------------------------------------
```

```
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----------------------------------------------------------------------
DR   EMBL; M30186; AAA25657.1; -.
DR   PIR; A35140; A35140.
KW   Hemolysis; Toxin; Outer membrane; Signal.
FT   SIGNAL         1     29
FT   CHAIN         30   1577       HEMOLYSIN.
SQ   SEQUENCE   1577 AA;  165869 MW;  175975E0C924B2D9 CRC64;

  Query Match             3.9%;  Score 183.5;  DB 1;  Length 1577;
  Best Local Similarity   20.3%;  Pred. No. 0.11;
  Matches  211;  Conservative  132;  Mismatches  371;  Indels  325;  Gaps   53;

Qy          3 NINVIKSN--------------IQAGLNSTKSGLKNLYLAIPKDYDPQKGGTLNDFIKAA 48
                |:|| |:|               |  |  |   :  :  :  |     |
Db        509 NLNVQKTNNDKTVTDNHVMWGGIGGGQNKNNNNQQQVSHATQLTAD-------------- 554

Qy         49 DELGIARLAEEPNHTETAKKSVDTVNQFLSLTQTGIAISATKLEKFLQKHSTNKLAKGLD 108
                 |  || : |   |  :       |:  ||  : |             | |:   :
Db        555 ---GQLLLAADNNVNITGSQVKGNQGAFVKTTQGDVVID------------NALS---E 595

Qy        109 SVENIDRKLGKASNVLSTLSSFLGTALAGIELDSLIKKGDAAPDALAKASIDLIN----E 164
               ::  || : | | |:  | |         :   |:      : :: | : :::
Db        596 TISKIDERTGTAFNI--TKSS---------HKNETNKQTSTGSELISDAQLTVVSGNDVN 644

Qy        165 IIGNLSQSTQTIEAFSSQLAKLGSTISQAKGFSNI--GNKLQNLNFSKTNLGLEIITGLL 222
               :||:| :|          |||      :  |  |:     ::  ::  ||:|    |||
Db        645 VIGSLIKSAD---------KLG---IHSLGDINVKSAQQVTKIDDEKTSLA---ITGHA 688

Qy        223 SGI-----SAGPALADKNASTGKKVAAGFELSNQVIGNVTKAISSYVLAQRVAAGLSTTG 277
                :      |||| :     :| |   :    | :|  |         : :       : | ||
Db        689 KEVEDKQYSAGFHITH---TTNKNTSTETEQANSTISGANVDLQANKDVTFAGSDLKTTA 745

Qy        278 AVAALITSSIMLAISPLAFMNAADK-------------FNHANALDEFAKQFRKFGYDGD 324
                 |::   ::          ||::  :|               |::   :|:   :   | ||
Db        746 GNASITGDNV-------AFVSTENKKQTDNTDTTISGGFSYTGGVDKVGSK-ADFQYDKQ 797

Qy        325 HLLAEY--QRGVGTIEASLTTISTALGAVSAGVSAAAVG----------------SAVGT 366
               |    |     ||  |    ||:          |   |
Db        798 HTQTEVTKNRGSQTEVAGDLTITANKDLLHEGASHHVEGRYQESGENIQHLAVNDSETSK 857

Qy        367 PIALLV---AGVTGLISGILEASKQAMFESVANRLQGKILEWEKQNGGQNYFDKGYDSRY 423
               :| |      ||    ||: :   |:|: :  |         |
Db        858 TDSLNVGIDVGVNLDYSGVTKPVKKAIEDGVNTTKPG----------------------- 894

Qy        424 AAYLANNLKFLSELNKELEAERVIAITQQRWDNNIGELAGI------TKLGERIKSGKAY 477
                  ||     ::| |:: |    ||          |:  |: :       ::|  || | :
Db        895 -----NN----TDLTKKVTARDAIA--------NLANLSNLETPNVGVEVG--IKGGGS- 934

Qy        478 ADAFEDGKKVEAGSNITLDAKTGIIDISNSNGKKTQALHFTSP----LLTAGTESRERLT 533
                : | : |       |       | |||  ::|      |    | :  |        ||| | : |
Db        935 QQSQTDSQAVSTSIN------AGKIDIDSNNKLHDQGTHYQSTQEGISLTANTHTSEATL 988

Qy        534 NGKYSYINKLKFGRVKNWQVTDGE--ASSKLDFSKVIQRVAETEGTDEIGLIVNAK---- 587
               :  : :: |  |               |:  |:|         :    | :|| :   ||
Db        989 DKHQTTFHETKGG---------GQIGVSTKTGSDITVAIKGEGQTTDNALMETKAKGSQF 1039

Qy        588 AGNDDIFVG---------------QGKMNIDGG---------DGHDRVFYSKDGGFGNIT 623
                 | || :                  :||  |: |               |   |   | |:
Db       1040 TSNGDISINVGENAHYEGAQFDAQKGKTVINAGGDLTLAQATDTHSE-SQSNVNGSANLK 1098

Qy        624 VDGT--SATEAGSYTVNRKVARGDIYHEVVKRQET-KVGKRTETIQYRDYELRKVGYGYQ 680
               |  |   |      | :            |  :|    | | | ||| |  :   :  ||      |:
Db       1099 VGTTPESKDYGGGFNA------GTTHHS--KEQTTAKVGTITGS---QGIELNA---GHN 1144

Qy        681 ST---DNLKSVEEVIGSQFNDV-----------FKGSKFNDIFHSGEGDDLLD-----GGA 722
                 |     :| | :::  :  | |.              ||:  :      :| |  : |      |
Db       1145 LTLQGTHLSSEQDIALNATNKVDLQSASSEHTEKGNNLSGGVQAGFGKKMTDDASSVNGL 1204

Qy        723 GDDRLFGGKGNDRLSGDEGDDLLDGGSGDDVLNGGAGNDVYIFRKGDGNDTLYDGTGNDK 782
               |  :   || :::      ||   :      ||:  :|    || |:
Db       1205 GSAQFAIGKQDEKSVSREGGTI--NNSGNLTIN---GNSVH------------------ 1240

Qy        783 LAFADANISDIMIERTKEGIIVKRNDHSGSINIPRWYITSNLQNYQSN-KTDHKIEQLIG 841
               |  |  |  |   :               || |  |       ::    :|::|  ||      ||
Db       1241 LQGAQVNSKDTQL-----------TSQSGDIEI----TSAQSTDYKNNWGTD------IG 1279

Qy        842 KDGSYITSDQIDKILQDKKDGTVITSQELKKLADENKSQK-------LSASDIASSLNKL 894
               :|     ::    | : ::|  | | :    | | :      ||        |      :   : ||
Db       1280 FNGK--KTNNTPKEVTEEKPATSIHNIGGKLLVNVEDQQKTSHQNATLETGTLTINSNK- 1336

Qy        895 VGSMALFG---TANSVSSN 910
                   : |  |    ||:||: |
Db       1337 --DLTLSGANVTADSVTGN 1353


RESULT 24
YL24_ANASP
ID   YL24_ANASP     STANDARD;      PRT;  1683 AA.
AC   Q8YV57;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Hypothetical WD-repeat protein all2124.
GN   ALL2124.
OS   Anabaena sp. (strain PCC 7120).
OC   Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
OX   NCBI_TaxID=103690;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21595285; PubMed=11759840;
RA   Kaneko T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,
RA   Watanabe A., Iriguchi M., Ishikawa A., Kawashima K., Kimura T.,
RA   Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
RA   Nakazaki N., Shimpo S., Sugimoto M., Takazawa M., Yamada M.,
RA   Yasuda M., Tabata S.;
RT   "Complete genomic sequence of the filamentous nitrogen-fixing
RT   cyanobacterium Anabaena sp. strain PCC 7120.";
RL   DNA Res. 8:205-213(2001).
CC   -!- SIMILARITY: Contains 13 WD repeats.
CC   -----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----------------------------------------------------------------------
DR   EMBL; AP003588; BAB73823.1; -.
DR   PIR; AF2071; AF2071.
DR   InterPro; IPR001680; WD40.
DR   Pfam; PF00400; WD40; 14.
DR   PRINTS; PR00320; GPROTEINBRPT.
DR   SMART; SM00320; WD40; 14.
DR   PROSITE; PS00678; WD_REPEATS_1; 3.
DR   PROSITE; PS50082; WD_REPEATS_2; 14.
DR   PROSITE; PS50294; WD_REPEATS_REGION; 1.
KW   Hypothetical protein; Repeat; WD repeat; Complete proteome.
FT   REPEAT     1070   1109       WD 1.
FT   REPEAT     1111   1150       WD 2.
FT   REPEAT     1153   1192       WD 3.
FT   REPEAT     1195   1236       WD 4.
FT   REPEAT     1280   1318       WD 5.
FT   REPEAT     1320   1359       WD 6.
FT   REPEAT     1361   1400       WD 7.
FT   REPEAT     1402   1441       WD 8.
FT   REPEAT     1444   1483       WD 9.
FT   REPEAT     1486   1525       WD 10.
FT   REPEAT     1528   1567       WD 11.
```

```
FT   REPEAT      1570   1609       WD 12.
FT   REPEAT      1611   1650       WD 13.
SQ   SEQUENCE    1683 AA;  185459 MW;  F7CB361FF54F7137 CRC64;

  Query Match             3.9%;  Score 180.5;  DB 1;  Length 1683;
  Best Local Similarity   19.6%;  Pred. No. 0.17;
  Matches  197;  Conservative  125;  Mismatches  354;  Indels  327;  Gaps   44;

Qy         31 KDYDPQKGGTLNDFIKAADELGIARLAEE--PNHTETAKKSVDTVNQFLSLTQTGIAISA 88
              : |  | ||          || | |       : |    |   | ||       |||||| |
Db        767 QSYQEQLGG-----IKGALERSCQGVYESLPPQLQECAKWI------FLSLTQLGEGTED 815

Qy         89 TKLEKFLQKHSTNKLAKGLDSVENIDRKLGKASNVLSTLSSFLGTALAGIELDSLIKKGD 148
              |:    :        |     ||                  |  ||:      |  | |:: |
Db        816 TRRRIYKSDLIVKKYPAGL---------------VEQTLNVLTNAKLVVINLEAEI---- 856

Qy        149 AAPDALAKASIDLINEIIGNLSQSTQTIEAFSSQLAKLGSTI--------SQAKGFSNIG 200
                 :| |:    :    |||     |:|      | :  ||:         : :     |
Db        857 ---EAQGKS----FSPTPPNLSTPFVTVEVAHEILIRHWSTLRWWLEENRDRLRKQRQIN 909

Qy        201 NKLQNLNFSKTNLGLEIITGLLSGISAGFALADKNASTGKKVAAGFEL----SNQVIGNV 256
              :  |                    ||  |  | |       |  ::|   ::    :::: :|
Db        910 HACQ-------------LWQQSGKQADFLL------QGARLAEAEDIYIYWTDELGADV 949

Qy        257 TKAISSYVL---------------AQRVAAGLSTTGAVAALITSSIMLAISPLAFMNAAD 301
               : | : :                |||     ||  |:        ::     ||:      :
Db        950 QEFIGACLAERKHQQLQAKNRLKQAQRAVVALSVLGIAS--------VSFGGLAYWQGRE 1001

Qy        302 -KFNHANALDEFAKQFRKFGYDGDHLLAEYQRGVGTIEASLTTISTALGAVSAGVSAA-A 359
               :|       ||:  ::              :||:  |                 | |:  | : ||
Db       1002 AQFREIAALNSSSQA---------NLLSHQQ---------------LAALIASLKAAQQ 1036

Qy        360 VGSAVGTPIALLVAGVTGLISGILEASKQAMFE-SVANRLQGK---ILEWEKQNGGQNYF 415
              |   :  |   |  :| || |         :||:||     |||:|     ::        ||
Db       1037 VNHVIAVPNNLKLATVTTL--------QQALFEMQERNRLEGHKDGVISISISRDGQTIA 1088

Qy        416 DKGYDSRYAAYLANNLKFLSELNKELEAERVIAI--------------TQQRWDNNIGEL 461
                 |      :  :   | :  ||   :|    ::                  | : |  : | |
Db       1089 SGSLDKTIKLWSRDGRLFRT-LNGHEDAVYSVSFSPDGQTIASGGSDKTIKLWQTSDGTL 1147

Qy        462 AGITKLGERIKSGKAYADAFEDGKKVEAGSN---------------ITLDA-KTGIIDIS 505
               |  |                    ||| :  : |:            :||        |:| :
Db       1148 LK-TITGHEQTVNNVYFS--PDGKNLASASSDHSIKLWDTTSGQLLMTLTGHSAGVITVR 1204

Qy        506 NSNGKKTQALHFTSPLLTAGTESRE-----------RLTNGKYSYINKLKFG------- 546
                 |   :|               :  ||:|: :                  :  ||     ::| | |
Db       1205 FSPDGQT---------IAAGSEDKTVKLWHRQDGKLLKTLNGHQDWVNSLSFSPDGKTLA 1255

Qy        547 ------RVKNWQVTDGEASSKL----------DFSKVIQRVAETEGTDEI---------- 580
                    :| |:: ||:     |             :||    : :|       : |
Db       1256 SASADKTIKLWRIADGKLVKTLKGHNDSVWDVNFSSDGKAIASASRDNTIKLWNRHGIEL 1315

Qy        581 -------------------GLIVNAK-------------------AGNDDI----FVGQG 598
                                 :| :|                      |||   :     |:  |
Db       1316 ETFTGHSGGVYAVNFLPDSNIIASASLDNTIRLWQRPLISPLEVLAGNSGVYAVSFLHDG 1375

Qy        599 K-MNIDGGDGHDRVFYSKDGGF-----GNITVDGTSATEAGSYTVNRKVARGDIYHEVVK 652
               :    | ||: ::::|:||              ||   :  |  |  |       :   |     | ||
Db       1376 SIIATAGADGNIQLWHSQDGSLLKTLPGNKAIYGISFTPQGDLIAS---ANAD---KTVK 1429

Qy        653 RQETKVGKRTETIQYRDYELRKVGY-------GYQSTDNLKSVEEVIGSQFNDVFKGSKF 705
               :  ||   :|:  |  |: ||  :         |||      :        :|        ||
Db       1430 IWRVRDGKALKTLIGHDNEVNKVNFSPDGKTLASASRDNTVKLWNVSDGKFKKTLKGHT- 1488

Qy        706 NDIFHSGEGDDLLDGGAGDDRLFGGKGNDRLSGDEGDDLLDGGSGDDVLNGGAGNDVYIF 765
              :::|       |              ||    || |:   |  |  ||: : :  | ||:
Db       1489 DEVFWVSFSPD-------------GKIIASASADKTIRLWDSFSGNLIKSLPAHNDLVYS 1535

Qy        766 RKGDGNDTLYDGTGNDKLAFADANISDIMIERTKEGIIVKRNDHSGSINIPRWYITSNLQ 825
               : : ::     |   ||                 : :  |: :|  ::    :    ||    |:
Db       1536 VNFNPDGSMLASTSADK----------TVKLWRSHDGHLL--HTFSGHSNVV--------- 1575

Qy        826 NYQSNKTDHKIEQLIGKDGSYITSDQIDKILQD-KKDGTVITS 867
                 | |:                  || || |    || ::   : || ::|:
Db       1576 -YSSS---------FSPDGRYIASASEDKTVKIWQIDGHLLTT 1608


RESULT 25
NODO_RHILV
ID   NODO_RHILV     STANDARD;      PRT;   284 AA.
AC   P15728;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-APR-1990 (Rel. 14, Last sequence update)
DT   01-MAR-1992 (Rel. 21, Last annotation update)
DE   Nodulation protein O.
GN   NODO OR NOLR.
OS   Rhizobium leguminosarum (biovar viciae).
OG   Plasmid sym pRL1JI.
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC   Rhizobiaceae; Rhizobium/Agrobacterium group; Rhizobium.
OX   NCBI_TaxID=387;
RN   [1]
RP   SEQUENCE FROM N.A., AND SEQUENCE OF 1-8.
RC   STRAIN=8401;
RX   MEDLINE=90151607; PubMed=2303029;
RA   Economou A., Hamilton W.D.O., Johnston A.W.B., Downie J.A.;
RT   "The Rhizobium nodulation gene nodO encodes a Ca2(+)-binding protein
RT   that is exported without N-terminal cleavage and is homologous to
RT   haemolysin and related proteins.";
RL   EMBO J. 9:349-354(1990).
RN   [2]
RP   SEQUENCE FROM N.A., AND SEQUENCE OF 4-18.
RX   MEDLINE=90078125; PubMed=2687250;
RA   de Maagd R.A., Wijfjes A.H.M., Spaink H.P., Ruiz-Sainz J.E.,
RA   Wijffelman C.A., Okker R.J.H., Lugtenberg B.J.J.;
RT   "nodO, a new nod gene of the Rhizobium leguminosarum biovar viciae
RT   sym plasmid pRL1JI, encodes a secreted protein.";
RL   J. Bacteriol. 171:6764-6770(1989).
CC   -!- FUNCTION: THE NODO PROTEIN MAY PLAY A ROLE IN NODULE DEVELOPMENT
CC       BY DIRECT INTERACTION WITH THE ROOT HAIR CELLS OR SOME OTHER PLANT
CC       SURFACE IN A CA(2+)-DEPENDENT MANNER.
CC   -!- SUBCELLULAR LOCATION: SECRETED. BY A MECHANISM THAT DOES NOT
CC       INVOLVE AN N-TERMINAL SIGNAL PEPTIDE.
CC   -!- INDUCTION: BY NARINGENIN (FLAVONOID).
CC   -!- SIMILARITY: TO HEMOLYSIN AND RELATED PROTEINS.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; X17285; CAA35178.1; -.
DR   EMBL; M29532; AAA26341.1; -.
DR   PIR; A43721; A43721.
DR   PIR; S08385; S08385.
DR   InterPro; IPR001343; Hemlysn_Ca_bind.
DR   Pfam; PF00353; hemolysinCabind; 5.
DR   PRINTS; PR00313; CABNDNGRPT.
DR   PROSITE; PS00330; HEMOLYSIN_CALCIUM; 1.
KW   Plasmid; Nodulation; Calcium-binding; Repeat.
FT   DOMAIN        2     73       8 X APPROXIMATE TANDEM REPEATS.
FT   DOMAIN       92    127       4 X APPROXIMATE TANDEM REPEATS.
FT   CA_BIND      92    131       POTENTIAL.
FT   DOMAIN      208    222       EXPORT SIGNAL (ASPARTIC ACID BOX)
FT                                (POTENTIAL).
SQ   SEQUENCE   284 AA;  30002 MW;  9CA41DCFCBDF8E15 CRC64;

  Query Match             3.8%;  Score 178.5;  DB 1;  Length 284;
  Best Local Similarity   23.2%;  Pred. No. 0.021;
  Matches   61;  Conservative   38;  Mismatches   99;  Indels   65;  Gaps    9;

Qy        684 NLKSVEE---VIGSQFNDVFKGSKFNDIFHSGEGDDLL---DG------GAGDDRLFGGK 731
              |:|  :     : ||  ||:  | | ||    :| |||  :     ||         | | |  :: ||
```

```
Db          2 NIKGSDNGSFIKGSPENDIIDGGKKNDWIDAGNGDDRIKAGDGQDSITAGPGHDIVWAGK 61

Qy        732 GNDRLSGDEGDDLL------------------DGGSGDDVLNGGAGNDVYIFRKGDGNDT 773
              |:|  :  | |||||                  | |||||  | |:|: :    ||| |
Db         62 GSDVIHADGGDDLLYSDASYPLYVTDPHRVIPHSGEGDDVLYAGPGSDILV--AGDGADV 119

Qy        774 LYDGTGNDKLAFADANISDIMIERT------------KEGIIVKRNDHSGSINIPRWYIT 821
              |  |   |   |       | |:  |            :: ::  |  |  |:
Db        120 LTGGDDGDAFVF---RFHDPMVGTTHCYTSVMDFDTKQDRFVLDAAPFGGDRNL----FD 172

Qy        822 SNLQNYQSN--------------KTDHKIEQLIGKDGSYITSDQIDKILQDKKDGTVITS 867
               :| |:                :  |   :: |  | : ::      :  :  | :|
Db        173 ANFINHSKGFPGEFVDTFYNGAAEGAHGEHVVVITDRGFASAAAAATAIDHEARGDIIVF 232

Qy        868 QELKKLADENKSQKLSASDIASS 890
               : | |  : ::   : : : |:
Db        233 HDQKTLGQDGETHGATLAYVDSA 255
```

RESULT 26
SLAP_CAUCR

```
ID   SLAP_CAUCR     STANDARD;      PRT;  1025 AA.
AC   P35828; Q46015; Q9RFI2;
DT   01-JUN-1994 (Rel. 29, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   S-layer protein (Paracrystalline surface layer protein).
GN   RSAA OR CC1007.
OS   Caulobacter crescentus.
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Caulobacterales;
OC   Caulobacteraceae; Caulobacter.
OX   NCBI_TaxID=155892;
RN   [1]
RP   SEQUENCE FROM N.A., AND SEQUENCE OF 1-21; 646-660 AND 1021-1025.
RC   STRAIN=ATCC 19089 / CB15;
RX   MEDLINE=93007489; PubMed=1393820;
RA   Gilchrist A., Fisher J.A., Smit J.K.;
RT   "Nucleotide sequence analysis of the gene encoding the Caulobacter
RT   crescentus paracrystalline surface layer protein.";
RL   Can. J. Microbiol. 38:193-202(1992).
RN   [2]
RP   REVISIONS TO 376; 636 AND 842-843.
RA   Awram P.;
RL   Submitted (OCT-1999) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=JS3001;
RA   Bingle W.H., Awram P.A., Nomellini J.F., Smit J.K.;
RT   "The secretion signal of C. crescentus S-layer protein is located in
RT   the C-terminal 82 amino acids of the molecule.";
RL   Submitted (OCT-1999) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 19089 / CB15;
RX   MEDLINE=21173698; PubMed=11259647;
RA   Nierman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,
RA   Eisen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,
RA   Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,
RA   DeBoy R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H.,
RA   Kolonay J.F., Smit J., Craven M.B., Khouri H., Shetty J., Berry K.,
RA   Utterback T., Tran K., Wolf A., Vamathevan J., Ermolaeva M., White O.,
RA   Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;
RT   "Complete genome sequence of Caulobacter crescentus.";
RL   Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).
RN   [5]
RP   SEQUENCE OF 1-313 FROM N.A., AND SEQUENCE OF 1-20.
RC   STRAIN=ATCC 19089 / CB15;
RX   MEDLINE=89008089; PubMed=3049545;
RA   Fisher J.A., Smit J.K., Agabian N.;
RT   "Transcriptional analysis of the major surface array gene of
RT   Caulobacter crescentus.";
RL   J. Bacteriol. 170:4706-4713(1988).
RN   [6]
RP   CHARACTERIZATION.
RC   STRAIN=ATCC 19089 / CB15;
RX   MEDLINE=98292737; PubMed=9620954;
RA   Awram P., Smit J.K.;
RT   "The Caulobacter crescentus paracrystalline S-layer protein is
RT   secreted by an ABC transporter (type I) secretion apparatus.";
RL   J. Bacteriol. 180:3062-3069(1998).
CC   -!- FUNCTION: THE S-LAYER IS A PARACRYSTALLINE MONO-LAYERED ASSEMBLY
CC       OF PROTEINS WHICH COAT THE SURFACE OF BACTERIA. PROBABLY ACTS AS A
CC       PHYSICAL BARRIER TO PARASITES AND LYTIC ENZYMES.
CC   -!- SUBCELLULAR LOCATION: CELL WALL. THIS BACTERIUM IS COVERED BY A S-
CC       LAYER WITH HEXAGONAL SYMMETRY. SECRETED BY AN ABC TRANSPORTER
CC       (TYPE I) SECRETION APPARATUS.
CC   -!- MISCELLANEOUS: THE CALCIUM BINDING OF THIS PROTEIN MAY MEDIATE
CC       SURFACE ATTACHMENT BY IONIC BRIDGING BETWEEN THE S-LAYER AND A
CC       SPECIFIC MEMBRANE-ASSOCIATED OLIGOSACCHARIDE.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; AF062345; AAC38665.2; -.
DR   EMBL; AF193063; AAF19365.1; -.
DR   EMBL; AE005779; AAK22991.1; ALT_INIT.
DR   PIR; A48995; A48995.
DR   HSSP; P22629; 1SWC.
DR   TIGR; CC1007; -.
DR   InterPro; IPR001343; Hemlysn_Ca_bind.
DR   Pfam; PF00353; hemolysinCabind; 3.
DR   PRINTS; PR00313; CABNDNGRPT.
KW   Cell wall; S-layer; Calcium-binding; Complete proteome.
FT   INIT_MET      0      0
SQ   SEQUENCE   1025 AA;  98001 MW;  AD7A326E1363D8AC CRC64;

  Query Match             3.8%;  Score 178.5;  DB 1;  Length 1025;
  Best Local Similarity   19.2%;  Pred. No. 0.11;
  Matches  169;  Conservative  116;  Mismatches  389;  Indels  205;  Gaps   36;

Qy          2 SNINVIKSNIQAGLNSTKSGLKNLYLAIPKDYDPQKGGTLNDFIKAADELGIARLAEEP- 60
              :||: : : ::|  |:  :   :: ||:        |   :    :  ||   ||   |
Db        166 ANIDYLTAFVRA--NTPFTAAADIDLAV-------KAALIGTILNAATVSGIGGYATATA 216

Qy         61 ---NHTETAKKSVDT---VNQFLSLTQTGIAISATKLEKFLQKHSTNKLAKGLDSVENID 114
                 |     |     ||| |  :|::             ||  |   | |::
Db        217 AMINDLSDGALSTDNAAGVNLFTAYPSSGVS-----------GSTLSLTTGTDTL---- 260

Qy        115 RKLGKASNVLSTLSSFLGTAL----------AGIELDSLIKKG--DAAPDALAKASIDLI 162
                |  |:|            | |         || :: : ::     |  |   :  :|: :
Db        261 --TGTANNDTFVAGEVAGAATLTVGDTLSGGAGTDVLNWVQAAAVTALPTGVTISGIETM 318

Qy        163 NEIIG-----NLSQSTQTIEAFSSQLAKLGSTISQAKGFSNIGNKLQNLNF-----SKTN 212
               |   |    ||  |      :  |  :: :     |::   |         |||      :  |
Db        319 NVTSGAAITLNTSSGVTGLTALNTNTSGAAQTVTAGAG--------QNLTATTAAQAANN 370

Qy        213 LGLE---IITGLLSGISAGFALADKNASTGKKVAAGFELSN-----------------Q 251
              : ::     :|    :|:::|          |::      |:       |:                 |
Db        371 VAVDGGANVTVASTGVTSGTTTVGANSAASGTVSVSVANSSTTTTGAIAVTGGTAVTVAQ 430

Qy        252 VIGNVTKAISSYVLAQRVAAGLSTTGAVAALITSSIMLAISPLAFMNAADKFNHANALDE 311
                ||       ::      |      | |:| ||      ::          :  |     :| |
Db        431 TAGNAVN--TTLTQADVTVTGNSSTTAVTVTQTAAATAGATVAGRVNGA-----VTITDS 483

Qy        312 FAKQFRKFGYDGDHLLAEYQRGVGTIEAS-LTTISTALGAVSAGVSAAAVGSAVGTPIA- 369
                |          |   :  |  ||::| |||::       | |:         |:  |||
Db        484 AAASATTAGKIATVTLGSF--GAATIDSSALTTVNLSGTGTSLGIGR---GALTATPTAN 538

Qy        370 LLVAGVTGL-ISGIL---EASKQAMFESVANRLQGKILEWEKQNGGQNYFDKGYDSRYAA 425
                |   | ||  :|  :     ||:      | ::                       |     | |:
Db        539 TLTLNVNGLTTTGAITDSEAAADDGFTTI------------------NIAGSTASSTIAS 580
```

```
Qy        426 YL---ANNLKFLSELNKELEAERVIAITQQRWDNNIGELAGITKLGERIKSGKAYADAFE 482
Db        581 LVAADATTLNISGDARVTITSHTAAALTGITVTNSVGATLGAELATGLVFTGGAGADSIL 640

Qy        483 DGKKVEAGSNITLDAKTGIIDISNS--------NGKKTQALHFTSPLLTAGTESRERLTN 534
Db        641 LGATTKA---IVMGAGDDTVTVSSATLGAGGSVNGGD------------GTDVLVANVN 684

Qy        535 GKYSYINKLKFGRVKNWQVTDGEAS--------SKLDFSKVIQRVAETEGTDEIGLIVNA 586
Db        685 GS-SFSADPAFGGFETLRVAGAAAQGSHNANGFTALQLGATAGATTFTNVAVNVGLTVLA 743

Qy        587 KAGNDDIFVGQGKMNIDGGDGHDRVFYSKDGGFGNITVDGTSATEAGSYTVNRKVARGDI 646
Db        744 AP------TGTTTVTLANATGTSDVF--------NLTLSSSAALAAGT------VALAGV 783

Qy        647 YHEVVKRQETKVGKRTETIQYRDYELRKV------GYGYQSTDNL-----KSVEEVIGS 694
Db        784 ETVNIAATDTNTTAHVDTLTLQATSAKSIVVTGNAGLNLTNTGNTAVTSFDASAVTGTGS 843

Qy        695 QFNDVFKGSKFNDIFHSGEGDDLLDGGAGDDRLFG-GKGNDRLSGDEGDDLLDGGSGDDV 753
Db        844 AVTFVSANTTVGEVV-------TIRGGAGADSLTGSATANDTIIGGAGADTLVYTGGTDT 896

Qy        754 LNGGAGNDVYIFRKGDGNDTLYDGTGNDKLAFADANISD 792
Db        897 FTGGTGADIF--------DINAIGTSTAFVTITDAAVGD 927
```

```
RESULT 27
SLAP_BACST
ID   SLAP_BACST     STANDARD;      PRT;  1228 AA.
AC   P35825;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   01-NOV-1995 (Rel. 32, Last annotation update)
DE   S-layer protein precursor (Surface layer protein).
GN   SBSA.
OS   Bacillus stearothermophilus.
OC   Bacteria; Firmicutes; Bacillales; Bacillaceae; Geobacillus.
OX   NCBI_TaxID=1422;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PV72;
RX   MEDLINE=94320770; PubMed=8045409;
RA   Kuen B., Sleytr U.B., Lubitz W.;
RT   "Sequence analysis of the sbsA gene encoding the 130-kDa
RT   surface-layer protein of Bacillus stearothermophilus strain PV72.";
RL   Gene 145:115-120(1994).
CC   -!- FUNCTION: THE S-LAYER IS A PARACRYSTALLINE MONO-LAYERED ASSEMBLY
CC       OF PROTEINS WHICH COAT THE SURFACE OF BACTERIA.
CC   -!- SUBCELLULAR LOCATION: CELL WALL. THIS BACTERIUM IS COVERED BY A
CC       S-LAYER WITH HEXAGONAL SYMMETRY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; X71092; CAA50409.1; -.
DR   EMBL; AX000218; CAB77069.1; -.
DR   PIR; I40468; I40468.
KW   Signal; Cell wall; S-layer.
FT   SIGNAL        1     30       POTENTIAL.
FT   CHAIN        31   1228       S-LAYER PROTEIN.
SQ   SEQUENCE   1228 AA;  131076 MW;  9F1383AD810C0B0F CRC64;

  Query Match             3.8%;  Score 175.5;  DB 1;  Length 1228;
  Best Local Similarity   19.4%;  Pred. No. 0.21;
  Matches  210;  Conservative  154;  Mismatches  355;  Indels  363;  Gaps   53;
```

```
Qy         15 LNSTKSGLKNLYLA-IPKDYDPQKGGTLNDFIKAADELGIARLAEEPNHTETAKKSVDTV 73
Db         86 VNKAGGAKKDAYLADLQKEYETY-------VFKANPKSGEARVA----------TYIDAY 128

Qy         74 NQFLSLTQTGIAISATKLEKFLQKHSTNKLAKGLDSVEN---------------IDRKLG 118
Db        129 NY-----------ATKLDEMRQELEAAVQAKDLEKAEQYYHKIPYEIKTRTVILDRVYG 176

Qy        119 -------------KASNVLSTLSSFLGTALAGIELDSLIKKGDAAPDALAKASIDLINEI 165
Db        177 KTTRDLLRSTFKAKAQELRDSLIYDITVAMKAREVQDAVKAGNLDK---AKAAVDQINQY 233

Qy        166 IGNLSQSTQTIEAFSSQLAKLGSTISQAKGFSNIGNKLQNLNFSKT-NLGLEIITGLLSG 224
Db        234 LPKVT------DAFKTELTEVAKKALDADE-AALTPKVESVSAINTQNKAVELTAVPVNG 286

Qy        225 -----ISAGFALADKNASTGKKVAAGFELSNQVIGNVTKAISSYVLAQRVAAGLSTTGAV 279
Db        287 TLKLQLSAAANEDTVNVNTVR--------IYKVDGNIPFALNT------ADVSLSTDGKT 332

Qy        280 AALITSSIMLAISPLAFMNAAD-KFNHANALDEFAKQFRK----FGYDGDHLLAEYQRGV 334
Db        333 ITVDAST--------PFENNTEYKVVVKGIKDKNGKEFKEDAFTFKLRNDAVVTQV---F 381

Qy        335 GTIEASLTTISTALGA----------------------------VSAGVSAAAVGSAVG 365
Db        382 GTNVTNNTSVNLAAGTFDTDDTLTVVFDKLLAPETVNSSNVTITDVETGKRIPVIASTSG 441

Qy        366 TPIALLV--AGVTG-----------LISGILEASKQAMFESVAN---------RLQGKIL 403
Db        442 STITITLKEALVTGKQYKLAINNVKTLTGYNAEAYELVFTANASAPTVATAPTTLGGTTL 501

Qy        404 E--------WEKQNGGQN-----YFDKGYDSRYA-----AYLANNLKFL---------SE 436
Db        502 STGSLTTNVWGKLAGGVNEAGTYYPGLQFTTTFATKLDESTLADNFVLVEKESGTVVASE 561

Qy        437 LNKELEAERVIAITQ-QRWDNNIGELAGITKLGERIKSGKAYADAFEDGKKVEAGSNITL 495
Db        562 LKYNADAKMVTLVPKADLKENTIYQI--------KIKKG------LKSDKGIELG---TV 604

Qy        496 DAKT-----------GIIDISNSNG----KKTQALHFTSPL------LTAGTESRERLTN 534
Db        605 NEKTYEFKTQDLTAPTVISVTSKNGDAGLKVTEAQEFTVKFSENLNTFNATTVSGSTITY 664

Qy        535 GKYSYINKLKFGRVKNWQVTDGEASSKLDFSKVI----QRVAETEGTDEIGLIVNAKAGN 590
Db        665 GQVAVVKA-------------GANLSALTASDIIPASVEAVTGQDGTYKVKVAAN----- 706

Qy        591 DDIFVGQGKMNIDGGDGHDRVFYSKDGGFGNITVDGTSATEAGSYTVNRKVARGDIYHEV 650
Db        707 ----------QLERNQGYKLVVFGK-----------GATAPVKDAANAN-TLATNYIYTFT 745

Qy        651 VKRQETKVGKRTETIQYRDYELRKVGYGYQSTDNLKSVEEVIGSQFNDVFKGSKFNDIFH 710
Db        746 TEGQDVTAPTVTKVFK---------------GDSLKDADAV--TTLTNVDAGQKFTIQF- 787

Qy        711 SGEGDDLLDGGAGDDRLFGGKGNDRLSGDEGDDLLDGGSGDDVLNGGAGNDVYIFRKGDG 770
Db        788 ----SEELKTSSGS--LVGGKVT-------VEKLTNNGWVD----AGTGTTVSVAPKTDA 830

Qy        771 ND-------TLYDGTGNDKLAFADANISDIMIERTKEGI------IVKRNDHSGSINIPR 817
Db        831 NGKVTAAVVTLTGLDNNDK----DAKLRLVVDKSSTDGIADVAGNVIKEKD--------- 877

Qy        818 WYITSNLQNYQS-NKTDHKIEQLIGKDG-----SYITSDQIDKILQDKKDGTVITSQEL- 870
Db        878 -----ILIRYNSWRHTVASVKAAADKDGQNASAAFPTSTAID-----------TTKSLL 920

Qy        871 -----KKLADENKSQKLSASDIASSLNKLVGSMALFGTANSVSSNALQPITQPTQGILAP 925
Db        921 VEFNETDLA-EVKPENIVVKDAAGN--------AVAGTVTALDGSTNKFVFTPSQELKAG 971
```

```
Qy        926 SV 927
              :|
Db        972 TV 973

RESULT 28
RBP2_PLAVB
ID   RBP2_PLAVB     STANDARD;      PRT;  1251 AA.
AC   Q00799;
DT   01-APR-1993 (Rel. 25, Created)
DT   01-APR-1993 (Rel. 25, Last sequence update)
DT   01-OCT-1996 (Rel. 34, Last annotation update)
DE   Reticulocyte binding protein 2 (Fragment).
GN   RBP2.
OS   Plasmodium vivax (strain Belem).
OC   Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX   NCBI_TaxID=31273;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92315338; PubMed=1617731;
RA   Galinski M.R., Medina C.C., Ingravallo P., Barnwell J.W.;
RT   "A reticulocyte-binding protein complex of Plasmodium vivax
RT   merozoites.";
RL   Cell 69:1213-1226(1992).
CC   -!- FUNCTION: INVOLVED IN RETICULOCYTE ADHESION. SPECIFICALLY BINDS TO
CC       HUMAN RETICULOCYTE CELLS.
CC   -!- SUBCELLULAR LOCATION: Membrane-bound (Probable).
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; M88098; AAA29744.1; -.
KW   Malaria; Receptor; Membrane.
FT   NON_TER       1      1
FT   NON_TER    1251   1251
SQ   SEQUENCE   1251 AA;  143741 MW;  54BA51C7404AC572 CRC64;

  Query Match             3.7%;  Score 173;  DB 1;  Length 1251;
  Best Local Similarity   19.3%;  Pred. No. 0.28;
  Matches  186;  Conservative  142;  Mismatches  341;  Indels  294;  Gaps   41;

Qy         28 AIPKDYDPQKGGTLNDFIKAADELGIARLAEEP-------------------------- 60
              :| | |: :  | || :   |  ||  |  :  ||
Db         18 SIEKAYE-KMGNTLKELEKMDDEKNIEKEVEEAQIQYKRIFIDHDVNLMNDEVEKSKIVM 76

Qy         61 NHTETAKKSVDTV-----------------NQFLSLTQTGIAISATKLEKFLQ----KH 98
                  |  || :| :                |: | ||     |   |:|:|:    |
Db         77 EKIELYKKEIDEIKQKTNEYKQGDTSNFYYTEQYNSATQ-----SKAKIEQFINIATTKK 131

Qy         99 STNKLAKGLDSVENIDRKLGKASNVLSTLSSFLGTALAGIELDSLIKKGDAAPDALAKAS 158
                 |:  :: :: :|:|  :: |  |:                  |:|:   : : : |
Db        132 GTSDTSQDINELESIKEEVHK--NL-----------------QLVKQESNSMEEMRKQI 171

Qy        159 IDLINEIIGNLSQSTQTIEAFSSQLAK-LGSTISQAKGF-SNIGNKLQNLNFSKTNLGLE 216
              : : : :| | |::             :|| : :     | ||   |    ||      :||: |:
Db        172 LSMKDLLILNNSET----------IAKEISNNTQNALGFRENAKTKL-----NKTDELLQ 216

Qy        217 IITGLLSGISA-----GFALADKNASTGKKVAAGFELSNQVIGNVTKAISSYVLAQRVAA 271
               :  ::    |        || |   |      :  |  | |: |  |     | ||:   :
Db        217 RVAAMIEEAKAHKNNIDIALEDAQIDT---EVSKIEQINREIMNKKDEIKSYLSEIKEYK 273

Qy        272 GLSTTGAVAALITSSIMLAISPLAFMNAADKFNHANALDEFAKQFRKFGYDGDHL--LAE 329
                 ||     |::|               ||         || ::|:      :    : |
Db        274 DKCTTE-----ISNS----------KRGKDKI-------EFLEKFKPNEESNSNKVNINE 311

Qy        330 YQRGVGTIEASLTTISTALGAVSAGVSAAAVGSAVGTPIALLVAGVTGLISGILEASK-- 387
                :  |  |  |  |   |                 | :  |    |  || | :  |:

Db        312 INENIRNSEQYLKDIEDAEKQAS-------------TKVELFHKHET-TISNIFKESEIL 357

Qy        388 --QAMFESVANRLQGKILEWEKQNGGQNYFDKGYDSRYAAYLANNLKFLS---------- 435
                :   :   |: :  : | |: |        ||:              || |:
Db        358 GVETKSQKKINKAEDIMKEIERHNSEIQTQVKGFQ--------ENLNKLNEPHNYDNAED 409

Qy        436 ELN--KELEAERVIAITQQRWDNNIGELAGITKLGERIKSGKAYADAFEDGKKVEAGSNI 493
              |||   |     |: :|      :    :|: |:  |': ||:|      |: |  :  :|::| |
Db        410 ELNNDKSTNAKVLIETNLESVKHNLSEITNIKQGGEKI-----YSKAKDIMQKIKATSEN 464

Qy        494 TLDAKTGIIDISNSNGKKTQALHFTSPLLTAGTESRERLTNGKYSYINKL-KFGRVKNWQ 552
              | :                                  :: |:: : : :|:| |  :   :|
Db        465 TAE-----------------------------KTLEKVKDDQSNYVNYLNQITTERNLI 494

Qy        553 VTDGEASSKLD--FSKVIQRVAETEGTDEIGLIVNAK--AGNDDIFVGQGKMNIDGGDGH 608
              ||:     : :|       : :    :  |::| |||  :      :  |   :  |  |:|:   |:
Db        495 VTEKNRLNGIDSTITNIEGALKESKGNYEIGFLEKLEEIGKNRKLKVDITKKSINSTVGN 554

Qy        609 DRVFYSKDGGFGNITVDGTSATEAGSYTVNRKV-----ARGDIYHEVVKRQETKVGKRTE 663
                 |  |    |              :    |   |: :         |:||:|     |  : |  :|
Db        555 ---FSSLFNNF----------DLNQYDFNKNINDYENKMGEIYNEF----EGSLNKISE 596

Qy        664 TIQ-----YRDYELRKV--GYGYQSTDNLKSVEEVIGSQFNDVFKGSKFNDIFHSGEGDD 716
               ::        ||   |       :     || : ||         || |    |  ||:   |   |
Db        597 NLRNASENTSDYNSAKTLRLEAQKEKVNLLNKEEEANKYLRDVKKVESFRFIFNMKESLD 656

Qy        717 LLDGGAGDDRLFGGKGNDRLSGDEGDDLLDGGSGDDVLNGGAGNDVYIFRKGDGNDTLYD 776
               ::      ::|                         .       :| |  ||     :
Db        657 KINEMIKKEQL-----------------------TVNEGHGNVKQLVE---------- 681

Qy        777 GTGNDKLAFADANISDIMIERTKEGIIVKRNDHSGSINIPRWYITSNLQNYQSNKTDHKI 836
                 |  |        : |:|||: :  | |  :      :::  |            | |:|       | :
Db        682 ---NIKELVDENNLSDILKQATGKNEEIQKITH------------STLKNKAKTILGH-V 725

Qy        837 EQLIGKDGSYITSD-QIDKILQDKKDGTVITSQELK-------KLADENKSQKLSASDIA 888
               :        |  ||  :   : ::|  | |    :  |:||||         |  || |:   :   |:
Db        726 DTSAKYVGIKITPELALTELLGDAK---LKTAQELKFESKNNVVLETENMSKNTNELDVH 782

Qy        889 SSL 891
               ::
Db        783 KNI 785


RESULT 29
ZAPA_PROMI
ID   ZAPA_PROMI     STANDARD;      PRT;   491 AA.
AC   Q11137;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Extracellular metalloprotease precursor (EC 3.4.24.-).
GN   ZAPA.
OS   Proteus mirabilis.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Proteus.
OX   NCBI_TaxID=584;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BB2000;
RX   MEDLINE=96011363; PubMed=7592325;
RA   Wassif C., Cheek D., Belas R.;
RT   "Molecular analysis of a metalloprotease from Proteus mirabilis.";
RL   J. Bacteriol. 177:5790-5798(1995).
CC   -!- FUNCTION: ONE OF THE VIRULENCE FACTORS PRODUCED DURING SWARMER
CC       CELL DIFFERENTIATION OF THE BACTERIA, WHICH SEEMS TO BE ASSOCIATED
CC       WITH PATHOGENESIS. THE PROTEASE ACTIVITY IS LIMITED TO IG A1,
CC       IG A2, AS WELL AS IG G DEGRADATION.
CC   -!- COFACTOR: Binds 1 zinc ion per subunit (By similarity).
CC   -!- ENZYME REGULATION: Ca(2+) increases protease activity.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- MISCELLANEOUS: THE GLY-RICH REPEATS MAY BE IMPORTANT IN THE
CC       EXTRACELLULAR SECRETION OF THIS METALLOPROTEASE.
CC   -!- SIMILARITY: Belongs to peptidase family M10B.
```

```
CC   -----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----------------------------------------------------------------------
DR   EMBL; U25950; AAA86729.1; ALT_INIT.
DR   HSSP; P23694; 1SAT.
DR   MEROPS; M10.057; -.
DR   InterPro; IPR001343; Hemlysn_Ca_bind.
DR   InterPro; IPR006026; NZn_MTpeptdse.
DR   InterPro; IPR006025; Zn_MTpeptdse.
DR   Pfam; PF00353; hemolysinCabind; 3.
DR   PRINTS; PR00313; CABNDNGRPT.
DR   SMART; SM00235; ZnMc; 1.
DR   PROSITE; PS00142; ZINC_PROTEASE; 1.
DR   PROSITE; PS00330; HEMOLYSIN_CALCIUM; FALSE_NEG.
KW   Hydrolase; Metalloprotease; Zinc; Zymogen; Repeat; Virulence; Calcium.
FT   PROPEP        1     16       POTENTIAL.
FT   CHAIN        17    491       EXTRACELLULAR METALLOPROTEASE.
FT   METAL       186    186       ZINC (CATALYTIC) (BY SIMILARITY).
FT   ACT_SITE    187    187       BY SIMILARITY.
FT   METAL       190    190       ZINC (CATALYTIC) (BY SIMILARITY).
FT   METAL       196    196       ZINC (CATALYTIC) (BY SIMILARITY).
FT   METAL       226    226       ZINC (CATALYTIC) (BY SIMILARITY).
FT   DOMAIN      343    393       4 X REPEATS, GLY-RICH.
FT   REPEAT      343    348       1.
FT   REPEAT      361    366       2.
FT   REPEAT      379    384       3.
FT   REPEAT      388    393       4.
SQ   SEQUENCE   491 AA;  54000 MW;  1E6EDEE3F6243A97 CRC64;

  Query Match              3.7%;  Score 170;  DB 1;  Length 491;
  Best Local Similarity   21.0%;  Pred. No. 0.12;
  Matches   98;  Conservative   75;  Mismatches  153;  Indels  140;  Gaps   22;

Qy        495 LDAKTGIIDISNSNGKKTQALHFTSPLL------TAGTE-SRERLT-NGKY--------- 537
              |   |: :||:   |      :::: :|       |||    :||  | ||||
Db          6 LKKAVGLSNISDLLDKSGIFYNYSTKVLPSFDYDTAGKHIAREDSTWNGKYVIGQPAEVT 65

Qy        538 ----SYINKL-KFGRVKNWQVTD---GEASSKLDFSKVIQRVAETEGTDEIGLIVNAKAG 589
                  :  |   :||    :  :           ||   :   ||      :|  :    ||
Db         66 YSFPKWEGKFNQFGNKNPYEFNELQKEHARKSLDAWSDIANIKFTEVA--VGNVDGMKAS 123

Qy        590 N--DDIFVGQGKMNIDGGDGHDRVF--------YSKD----GGFGNITVDGTSATEAGSY 635
               :    ||  |     ||   :|  :  :           | ||       |  :    |  :  |  |:|
Db        124 DVKTDITFG----NIYDPNGTFQAYATLPNTYAYGKDLSGQAWFSDYHYAGNTTPELGNY 179

Qy        636 TVNRKVARGDIYHEV---------------------VKRQETKVGKRTETIQYRD----- 669
                     |  |  ||:                      :|     :  ::    : |  |
Db        180 ------GRLTIIHEIGHTLGLMHPGDYNAGQNVPGYLKSDYAEDSRQYTVMSYWDEYETG 233

Qy        670 -------------YELRKVGYGYQSTDNLKSVEEVIGSQFN---DVFKGSKFND-----I 708
                           ::: :  | |  :    :: ::|  |    |     |  :  :  ||      :
Db        234 AHFQGAYAGAPLLHDISAMQYLYGANTTTRTGDDVYGFNSNTGIDYYTATDSNDKLIFSV 293

Qy        709 FHSGEG----------DDLLDGGAGD------------------DRLFGGKGNDRLSGD 739
              : ||            | |:|  ||:                       :    || ||| : |:
Db        294 WDSGGNDTFDFSGFYQDQLIDLRAGNFSDVGGLQKNVSIAQNVTIENAIGGFGNDIIHGN 353

Qy        740 EGDDLLDGGSGDDVLNGGAGND-VYIFRKGDGNDTLYDGTGNDKLAFADANISDIMIERT 798
              |: | ||  |||:: |  :||: :|     | |  |||: |||::        :   :  |:| ::
Db        354 DADNTLIGGEGDDIIYGHSGNNTIY---GGRGQDTLHGGTGSNTFIYKE--IADSLVTAA 408

Qy        799 KEGIIVKRNDHSGSINIPRWYITSNLQNYQSNKTDHKIEQLIGKDG 844
               : :    |            |  :   ::: :|:   |:|   : ::    |    |
Db        409 DKIMDFK-------TGIDKIDLSTLIQDTFSSKILNFVDNFTGNAG 447
```

RESULT 30

```
OMPA_RICRI
ID   OMPA_RICRI     STANDARD;      PRT;  2249 AA.
AC   P15921;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-APR-1990 (Rel. 14, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Outer membrane protein A precursor (190 kDa antigen) (Cell surface
DE   antigen) (rOmpA) (rOmp A).
GN   OMPA.
OS   Rickettsia rickettsii.
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC   Rickettsiaceae; Rickettsieae; Rickettsia.
OX   NCBI_TaxID=783;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=R;
RX   MEDLINE=90354033; PubMed=2117568;
RA   Anderson B.E., McDonald G.A., Jones D.C., Regnery R.L.;
RT   "A protective protein antigen of Rickettsia rickettsii has tandemly
RT   repeated, near-identical sequences.";
RL   Infect. Immun. 58:2760-2769(1990).
CC   -!- FUNCTION: ELICITS PROTECTIVE IMMUNITY.
CC   -!- SUBCELLULAR LOCATION: CELL WALL. THIS BACTERIUM IS COVERED BY A
CC       S-LAYER WITH HEXAGONAL SYMMETRY.
CC   -!- PTM: GLYCOSYLATED (PROBABLE).
CC   -!- SIMILARITY: BELONGS TO THE RICKETTSIAE OMPA/OMPB FAMILY.
CC   -----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----------------------------------------------------------------------
DR   EMBL; M31227; AAA26380.1; -.
DR   PIR; A41477; A41477.
DR   InterPro; IPR006315; Autotransport.
DR   InterPro; IPR005546; Autotransporter.
DR   Pfam; PF03797; Autotransporter; 1.
DR   TIGRFAMs; TIGR01414; autotrans_barl; 3.
KW   Antigen; Repeat; Signal; Cell wall; S-layer; Glycoprotein.
FT   SIGNAL        1     28       POTENTIAL.
FT   CHAIN        29   2249       OUTER MEMBRANE PROTEIN A.
FT   DOMAIN      212   1180       13 X APPROXIMATE TANDEM REPEATS.
FT   REPEAT      212    286       A (TYPE I).
FT   REPEAT      287    358       B (TYPE II).
FT   REPEAT      359    430       C (TYPE II).
FT   REPEAT      431    505       D (TYPE I).
FT   REPEAT      506    577       E (TYPE II).
FT   REPEAT      578    652       F (TYPE I).
FT   REPEAT      653    724       G (TYPE II).
FT   REPEAT      725    799       H (TYPE I).
FT   REPEAT      800    874       I (TYPE I).
FT   REPEAT      875    949       J (TYPE I).
FT   REPEAT      950   1021       K (TYPE II).
FT   REPEAT     1022   1093       L (TYPE II).
FT   REPEAT     1094   1165       M (TYPE II).
FT   REPEAT     1166   1180       TYPE I (INCOMPLETE).
SQ   SEQUENCE   2249 AA;  224333 MW;  A9D6646C089DF087 CRC64;

  Query Match              3.6%;  Score 169.5;  DB 1;  Length 2249;
  Best Local Similarity   20.3%;  Pred. No. 0.92;
  Matches  190;  Conservative  117;  Mismatches  338;  Indels  291;  Gaps   46;

Qy          1 MSNINVIKS-NIQAGLNSTKSGLKNLYLAIPKDYDPQKGGTLN--------DFIK----- 46
              : | | : : |: |||   : |:                           | |:|          |
Db        359 VGNTNALATVNVGAGLLQVQGGV-------------VKANTINLTDNASAVTFTNPVVVT 405

Qy         47 -AADELGIAR---LAEEPNHTETA----KKSVDTVNQFLSLTQTGIAI---SATKL---E 92
               | | |  |        :      | | |        :: |||          | |:   : |||
Db        406 GAIDNTGNANNGIVTFTGNSTVTGDIGNTNALATVNVGAGTATLGGAVIKATTTKLTNAA 465
```

```
Qy         93 KFLQKHSTNKLAKG-LDS--------VENIDRKLGKA------SNVLSTLSSFLGTALAG 137
Db        466 SVLTLTNANAVLTGAIDNTTGGDNVGVLNLNGALSQVTGNIGNTNSLATISVGAGTATLG 525

Qy        138 --IELDSLIKKGDAAPDALAKASIDLINEII--------GNLSQSTQTIEAFSSQLAKLG 187
Db        526 GAVIKATTTKLTDAA------SAVKFTNPVVVTGAIDNTGNANNGIVTFTGNSTVTGDIG 579

Qy        188 -----STISQAKGFSNIGNKLQNLNFSKTNLGLEIITGLLSGISAGFALADKNASTGKKV 242
Db        580 NTNSLATISVGAGTATLGGAVIKATTTKLTNAASVLT--LTNANAVLTGAIDNTTGGDNV 637

Qy        243 AAGFELS---NQVIGNVTKAISSYVLAQRVAAGLSTT-GAVAALITSSIMLAISPLAFMN 298
Db        638 GV-LNLNGALSQVTGDIGNTNSLATIS--VGAGTATLGGAVIKATTTKITNAVSAVKFTN 694

Qy        299 ------AADKFNHANALDEFAKQFRKFGYDGDHLLAEYQRGVGTIEASLTTI-----STA 347
Db        695 PVVVTGAIDSTGNAN-----------------------NGIVTFTGNSTVTGDIGNTNA 730

Qy        348 LGAVSAGVSAAAVGSAVGTPIALLVAGVTGLISGILEASKQAMFESVANRLQGKILEWEK 407
Db        731 LATVNVGAGTATLGGAV---IKATTTKLTNAASVLTLTNANAV-------LTGAI---DN 777

Qy        408 QNGGQNYFDKGYDSRYAAYLANNLKFLSELNKELEAERVIAITQQRWDNNIGELAGITKL 467
Db        778 TTGGDN-----------VGVLNLNGALSQVTGDI--------------GNTNSLATIS-- 810

Qy        468 GERIKSGKAYADAFEDGKKVEAGSNITLDAKTGIIDISNSNGKKTQALHFTSPLLTAGTE 527
Db        811 ---VGAGTATL-----GGAVIKATTTKLTNAASVLTLTNANAVLTGAVDNT----TGGDN 858

Qy        528 SRERLTNGKYSYI-------NKLKFGRVKNWQVTDGEASSKLDFSKVIQRVAETEGTDEI 580
Db        859 VGVLNLNGALSQVTGDIGNTNSLATISVGAGTATLGGAVIKATTTKLTNAASVLTLTNAN 918

Qy        581 GLIVNA---KAGNDDIFVGQGKMNIDGGDGHDRVFYSKDGGFGNITVDGTSATEAGSYTV 637
Db        919 AVLTGAIDNTTGGDNV----GVLNLNG------ALSQVTGDIGNTNSLATISVGAGTATL 968

Qy        638 NRKVARGDIYHEVVKRQETKVGKRTETIQYRDYELRKVGYGYQSTDNLKSVEEVIGSQFN 697
Db        969 G---------GAVIKATTTKL-----------------------TDAASAV--------- 987

Qy        698 DVFKGSKF-NDIFHSGEGDDLLDGGAGDDRLFGGKGNDRLSGDEGD-------------- 742
Db        988 ------KFTNPVVVTGAIDNT---GNANNGIVTFTGNSTVTGNVGNTNALATVNVGAGLL 1038

Qy        743 ------------DLLDGGSGDDVLN-----------GGAGNDVYIFRKGDGNDTLYDGTG 779
Db       1039 QVQGGVVKANTINLTDNASAVTFTNPVVVTGAIDNTGNANNGIVTF---TGNSTVTGNVG 1095

Qy        780 NDKLAFADANISDIMIERTKEGIIVKRNDHSGSINI 815
Db       1096 NTN-ALATVNVGAGLLQ--VQGGVVKAN----TINL 1124
```

RESULT 31
PMPD_CHLMU

```
ID   PMPD_CHLMU     STANDARD;      PRT;  1520 AA.
AC   Q9PLB0;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Probable outer membrane protein pmpD precursor (Polymorphic membrane
DE   protein D).
GN   PMPD OR TC0197.
OS   Chlamydia muridarum.
OC   Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
OX   NCBI_TaxID=83560;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MoPn / Nigg;
RX   MEDLINE=20150255; PubMed=10684935;
RA   Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
RA   White O., Hickey E.K., Peterson J., Utterback T., Berry K., Bass S.,
RA   Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,
RA   Gwinn M., Nelson W., DeBoy R., Kolonay J., McClarty G., Salzberg S.L.,
RA   Eisen J., Fraser C.M.;
RT   "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia
RT   pneumoniae AR39.";
RL   Nucleic Acids Res. 28:1397-1406(2000).
CC   -!- SUBCELLULAR LOCATION: CELL WALL SURFACE (ELEMENTARY BODIES)
CC       (POTENTIAL).
CC   -!- SIMILARITY: BELONGS TO THE PMP OUTER MEMBRANE PROTEIN FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; AE002286; AAF39070.1; -.
DR   PIR; A81731; A81731.
DR   TIGR; TC0197; -.
DR   InterPro; IPR005546; Autotransporter.
DR   InterPro; IPR003368; Chlamydia_PMP.
DR   Pfam; PF03797; Autotransporter; 1.
DR   Pfam; PF02415; DUF145; 1.
DR   TIGRFAMs; TIGR01376; POMP_repeat; 15.
KW   Outer membrane; Signal; Multigene family; Complete proteome.
FT   SIGNAL        1     20       POTENTIAL.
FT   CHAIN        21   1520       PROBABLE OUTER MEMBRANE PROTEIN PMPD.
SQ   SEQUENCE   1520 AA;  162102 MW;  68814405AC79585F CRC64;

  Query Match             3.6%;  Score 169;  DB 1;  Length 1520;
  Best Local Similarity   19.1%;  Pred. No. 0.58;
  Matches  222;  Conservative  143;  Mismatches  386;  Indels  412;  Gaps   58;

Qy         10 NIQAGLNSTKSGLKN---LYLAIPKDYD------PQKGGTLNDFIKAADELGIARLAEEP 60
Db         17 SVVAAILASMSGLSNCSDLY-AVGSSADHPAYLIPQ-AGLLLDHIK---DIFIG------ 65

Qy         61 NHTETAKKSVDTVNQFLSLTQTGIAISATKLEKFLQK--HS------TNKLAKGLDSVEN 112
Db         66 -----PKDSQDKGQYKLIIGEAG-SFQDSNAETLPQKVEHSTLFSVTTPIIVQGIDQQDQ 119

Qy        113 IDRK------LGKASNVLSTLSSFLGTALAGIELDSLIKKGDAAPD-------ALAKASI 159
Db        120 VSSQGLVCNFSGDHSEEIFERESFLGIAFLG----NGSKDGITLTDIKSSLSGAALYSSD 175

Qy        160 DLINE-IIGNLSQSTQTIEAFSSQLAKLGSTISQAKGFSNIGNKLQNLNFSKTNLGLEII 218
Db        176 DLIFERIKGDIELSS------CSSLERGGACSAQ----SILIHDCQGLTVKHCAAGVNV- 224

Qy        219 TGLLSGISA-------GFALADKNASTGKK---VAAGFELSNQVIGNVTKAISSYVLAQR 268
Db        225 ----EGVSASDHLGFGGGAFSTTSSLSGEKSLYMPAG--------DIVVATCDGPVCFEG 272

Qy        269 VAAGLSTTGAVAA------------------LITSSIMLAISPLAFMNAAD---KFNHA 306
Db        273 NSAQLANGGAIAASGKVLFVANEKKISFTDNQALSGGAISASSISFQNCAELVFKSNLA 332

Qy        307 NALDEFAKQFRKFGYDG-----------DHLLAEYQRGV-----GTIEASLTTISTALG 349
Db        333 KGVKD------KCSLGGGALASLESVVLKDNLGITYEKNQSYSEGGAIFGKDCEIFENRG 386

Qy        350 AVSAGVSAAAVGSAVGTPIALLVAGVTGLISGI-LEASKQAMFESVANRLQGKILEWEKQ 408
Db        387 PVVFRDNTAALGG--GAILAQQTVAICGNKSGISFEGSKSSF----------------- 426

Qy        409 NGGQNYFDKGYDSRYAAYLANNLKFLSELNKELEAERVIAITQQRWDNNIGELAGITKLG 468
Db        427 -GG------------AIACGN--FSSENNSS-------ALGSIDISNNLGDISFLRTLC 463
```

```
Qy        469 ERIKSGKAYADAFEDGKKVEAGSNITLDAKTGIIDISN-------SNGKKTQALHFTSPL 521

Db        464 TTSDLGQT---DYQGGGALFA-ENISLSENAGAITFKDNIVKTFASNGKMLGG----GAI 515

Qy        522 LTAGTESRERLTNGKYSYINKLKFGRVKNWQVTDGEASSKLDFSKVIQRVAETEGTD--- 578

Db        516 LASGNVLISK-NSGEISFV-----GNARAPQAIPTRSSDELSFG--AQLTQTTSGCSGGG 567

Qy        579 -----EIGLIVNAKAGNDDIFVGQGKMNIDGGDGHDRVFYSKDGGFGNITVDGTSATEAG 633

Db        568 ALFGKEVAIVQNATVVFEQNRLQCGEQETHGGGG---AVY----GMESASIIGNSFVRFG 620

Qy        634 -SYTVNRKVARGDIYHEVVKRQE-TKV--GKRTET-----IQYRDYELRKVGYGY----- 679

Db        621 NNYAVGNQISGGALLSKKVRLAENTRVDFSRNIATFCGGAVQVSDGSCELINNGYVLFRD 680

Qy        680 -------------------------QSTDN---------------------------- 684

Db        681 NRGQTFGGAISCLKGDVIISGNKDRVEFRDNIVTRPYFEENEEKVETADINSDKQEAEER 740

Qy        685 --LKSVEE---------------------------------------------------- 690

Db        741 SLLENIEQSFITATNQTFFLEEEKLPSEAFISAEELSKRRECAGGAIFAKRVYITDNKEP 800

Qy        691 -VIGSQFNDVFKGSKFND--------------IFHSGEGDDLLDGGAG--DDRL--FGG 730

Db        801 ILFSHNFSDVYGGAIFTGSLQETDKQDVVTPEVVISGNDGDVIFSGNAAKHDKHLPDTGG 860

Qy        731 ------------------KGNDRLSG------DEGDDLLDGGSGDDVLNGGA-----GN 760

Db        861 GAICTQNLTISQNNGNVLFLNNFACSGGAVRIEDHGEVLLEAFGGDIIFNGNSSFRAQGS 920

Qy        761 DVYIFRKGDGNDTLYDGTGNDKLAFADANISDIMIERTKEGIIV----KRNDHSGSI--- 813

Db        921 DAIYFAGKDSRIKALNATEGHAIVFQDALVFENIEERKSSGLLVINSQENEGYTGSVRFL 980

Qy        814 ----NIPRWYITSNLQNYQSNKTDHKIEQLIGKDGSYITSDQIDKILQDKKDGTVITSQE 869

Db        981 GSESKVPQWI----------HVQQGGLELL---HGAILCSYGVK---QDPRAKIVLSAGS 1024

Qy        870 LKKLADENKSQKLSASDIASSLN 892

Db       1025 KLKILDSEQENNAEIGDLEDSVN 1047
```

RESULT 32
OMPB_RICJA
ID   OMPB_RICJA   STANDARD;   PRT; 1656 AA.
AC   O06653;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Outer membrane protein B precursor (168 kDa surface-layer protein)
DE   (Surface protein antigen) (Cell surface antigen 5) (Sca5) (rOmpB)
DE   (rOmp B) [Contains: 120 kDa surface-exposed protein (Surface protein
DE   antigen) (120 kDa outer membrane protein ompB); 32 kDa beta peptide].
GN   OMPB.
OS   Rickettsia japonica.
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC   Rickettsiaceae; Rickettsieae; Rickettsia.
OX   NCBI_TaxID=35790;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=YH;
RA   Uchiyama T.;
RT   "Sequencing of the gene encoding the protein rOmp B of Rickettsia
RT   japonica.";
RL   Submitted (MAY-1997) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: THE 120 KDA SURFACE-EXPOSED PROTEIN IS A MAJOR
CC   STRUCTURAL PROTEIN WHICH MAY PLAY A ROLE AS A RICKETTSIAL
CC   VIRULENCE FACTOR AND/OR IMMUNOGEN DURING INFECTION (BY
CC   SIMILARITY).
CC   -!- FUNCTION: THE 32 KDA BETA PEPTIDE MAY SERVE AS A MEMBRANE ANCHOR
CC   (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: CELL WALL. THIS BACTERIUM IS COVERED BY A S-
CC   LAYER WITH HEXAGONAL SYMMETRY.
CC   -!- SIMILARITY: BELONGS TO THE RICKETTSIAE OMPA/OMPB FAMILY.
CC   --

CC   --
DR   EMBL; AB003681; BAA20138.1; -.
DR   InterPro; IPR006315; Autotransport.
DR   InterPro; IPR005546; Autotransporter.
DR   Pfam; PF03797; Autotransporter; 1.
DR   TIGRFAMs; TIGR01414; autotrans_barl; 2.
KW   Antigen; S-layer; Cell wall.
FT   CHAIN   1   1338   120 kDa SURFACE-EXPOSED PROTEIN.
FT   CHAIN   1339   1656   32 kDa BETA PEPTIDE.
FT   DOMAIN   528   533   POLY-GLY.
SQ   SEQUENCE   1656 AA; 168097 MW; 3132A69C9DD5999F CRC64;

Query Match 3.6%; Score 167; DB 1; Length 1656;
Best Local Similarity 19.7%; Pred. No. 0.82;
Matches 191; Conservative 133; Mismatches 355; Indels 290; Gaps 48;

```
Qy         56 LAEEPNHTETAKKSVDTVNQFL-SLTQTGI--AISATKLEKFL---------QKHSTNKL 103

Db          1 MAQKPN-------------FLKKLISAGLVTASTATIVASFAGSAMGAAIQQNRTTNGV 46

Qy        104 AKGLDSVENIDRKLGKASNVLSTLSSFLGTALAGIELDS---------LIKKGDAAPDAL 154

Db         47 ATTVDGV-GFDQTVALANVAVAPNAVITANANNGINLNTPAGSFNGLFLSNANNLAVTVS 105

Qy        155 AKASIDLINEIIGNLSQSTQTIEAFSSQLAKLGSTIS-QAKGFSNI----GNKLQNLNFS 209

Db        106 EDTTLGFINNAANNANRFNLTLDA--------GKTLTITGQGITNVQSAATHNAQNI-VA 156

Qy        210 KTNLGLEIITGLLSGISAGFALADKNASTGKKVAAGFELSNQVIGNVTKAISSYVLAQRV 269

Db        157 KFNGGAAIANNDLSGLGT----IDFGAAASTLV---FDLANP-----TTQKAPLILADN- 203

Qy        270 AAGLSTTGAVAALITSSIMLAISPLAF-----MNAAD----KFN----HANAL------- 309

Db        204 --ALIVNGANGTLNVTNGFIQVSDKSFATVKAINIGDGQGFMFNTNATNANALNLQAGGT 261

Qy        310 -----------------DEFAKQFRKFGYDGDHLLAEYQRGVGTIEASLTTISTALGAV 351

Db        262 TINFNGTDGTGRLVLLSKNGAATDFNVTGSLGGNL-------KGIIELNTVAINGQLIA- 313

Qy        352 SAGVSAAAVGSAVGTPIALLVAGVTGLISGILEASKQAMFESVANRLQGK--ILEWEKQN 409

Db        314 NAGPANAVIGTNNG----------AGRAAGFVVSVDNGKAATIDGQVYAKDMVIQSANAN 363

Qy        410 GGQNY---FDKGYDSRYAAYLANNLKFLSELNK-------ELEAERVIAITQQRWD---- 455

Db        364 GQVNFRHIVDVGIDGTTAFKTAASIVAITQNSNFGTTDFGNLAAQVTVPDTMTLTGNFTG 423

Qy        456 --NNIGELAGITKLGERIKSGKAYADAFEDGKKVEAGSNIT-LDAK-TGIIDISNS---- 507

Db        424 DANNPGNTAGVITFAANGTLASASADA-----NVAVTNNITAIEASGVGVVQLSGTHTAE 478

Qy        508 -----------------NGKKTQALHFTSPL---------------------------- 521

Db        479 LRLGNAGSVFKLADGTVINGKVNQTVLVGGVLAAGAITLDGSATITGDIGNGGGGAALQS 538

Qy        522 LTAGTESRERLTNGKYSYINKLKFGRVKNWQVTDGEASSKLDFSKVIQRVAETEGTDEIG 581

Db        539 ITLANDATKTLTLGGANIIS--ANGGTINFQANGGTIKLTSTQNNIVVDCDLAIATDQTG 596
```

```
Qy        582 LIVNAKAGNDDIFVGQGKMNIDGGD----GHDRVFYSK---DGGFGNITV-------DGT 627

Db        597 VVDASSLTNAQTLTISGTIGIIGANNTTLGQFNIGSSKTTLNG--GNVAINELVIGNNGS 654

Qy        628 SATEAGSYTVNR---KVARGDIYHEVVKRQETKVGKRTETIQYRDYELRKVGYGYQSTDN 684

Db        655 VQFAHNTYLITRTTNAAGQGKIIFNPVVNNNTTLAAGT--------------------N 693

Qy        685 LKSVEEVIGSQFNDVFKGSKFNDIFHSGEGDDL-------LDGGAGDDRLFGGKGNDRLS 737

Db        694 LGSAANPL-AEINFGSKGARADTVLNVGEGVNLYATNITTTDANVG-SFVFNAGGKNIVS 751

Qy        738 GDEGDD--------LLDGGSGDDVLNGGAGNDVYIFRKGDGNDTLYDGTGNDKLAFADAN 789

Db        752 GTVGGQQGNKFNTVALDNGTTVKFL----GNATF-----NGNTTI---AANSTLQISGNY 799

Qy        790 ISDIMIERTKEGIIVKRNDHSGSINIPRWYITSNLQNYQSNKTDHKIEQL---------- 839

Db        800 TADFIASADGTGIVEFVN--TGPIN-----VTLNKQAVPVN----ALKQITVSGPGNVVV 848

Qy        840 --IGKDGSY 846

Db        849 NEIGNAGNY 857
```

```
RESULT 33
PRTC_ERWCH
ID   PRTC_ERWCH     STANDARD;      PRT;   479 AA.
AC   P16317;
DT   01-AUG-1990 (Rel. 15, Created)
DT   01-NOV-1991 (Rel. 20, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Secreted protease C precursor (EC 3.4.24.-) (ProC).
GN   PRTC.
OS   Erwinia chrysanthemi.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Pectobacterium.
OX   NCBI_TaxID=556;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=B374;
RX   MEDLINE=91009140; PubMed=2211614;
RA   Delepelaire P., Wandersman C.;
RT   "Protein secretion in Gram-negative bacteria. The extracellular
RT   metalloprotease B from Erwinia chrysanthemi contains a C-terminal
RT   secretion signal analogous to that of Escherichia coli alpha-
RT   hemolysin.";
RL   J. Biol. Chem. 265:17118-17125(1990).
RN   [2]
RP   SEQUENCE OF 1-59 FROM N.A., AND PARTIAL SEQUENCE.
RC   STRAIN=B374;
RX   MEDLINE=89255387; PubMed=2722818;
RA   Delepelaire P., Wandersman C.;
RT   "Protease secretion by Erwinia chrysanthemi. Proteases B and C are
RT   synthesized and secreted as zymogens without a signal peptide.";
RL   J. Biol. Chem. 264:9083-9089(1989).
CC   -!- COFACTOR: Binds 1 zinc ion per subunit. Magnesium and calcium
CC       increase protease activity significantly above the control value.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- DOMAIN: THE GLY-RICH REPEATS MAY BE IMPORTANT IN THE EXTRACELLULAR
CC       SECRETION OF THIS METALLOPROTEASE.
CC   -!- SIMILARITY: Belongs to peptidase family M10B.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; M59229; AAA24860.1; -.
DR   EMBL; J04736; AAA24862.1; -.
DR   EMBL; M60395; AAA63638.1; -.
DR   PIR; A38307; A38307.
DR   PDB; 1GO7; 17-OCT-02.
DR   PDB; 1GO8; 17-OCT-02.
DR   PDB; 1K7G; 19-OCT-02.
DR   PDB; 1K7Q; 20-OCT-02.
DR   MEROPS; M10.054; -.
DR   InterPro; IPR001343; Hemlysn_Ca_bind.
DR   InterPro; IPR006026; NZn_MTpeptdse.
DR   InterPro; IPR006025; Zn_MTpeptdse.
DR   Pfam; PF00353; hemolysinCabind; 3.
DR   SMART; SM00235; ZnMc; 1.
DR   PROSITE; PS00142; ZINC_PROTEASE; 1.
DR   PROSITE; PS00330; HEMOLYSIN_CALCIUM; 1.
KW   Hydrolase; Metalloprotease; Zinc; Zymogen; Calcium; Repeat;
KW   3D-structure.
FT   PROPEP        1     17
FT   CHAIN        18    479       SECRETED PROTEASE C.
FT   METAL       188    188       ZINC (CATALYTIC) (BY SIMILARITY).
FT   ACT_SITE    189    189       BY SIMILARITY.
FT   METAL       192    192       ZINC (CATALYTIC) (BY SIMILARITY).
FT   METAL       228    228       ZINC (CATALYTIC) (BY SIMILARITY).
FT   DOMAIN      345    386       5 X REPEATS, GLY-RICH.
FT   REPEAT      345    350       1.
FT   REPEAT      354    359       2.
FT   REPEAT      363    368       3.
FT   REPEAT      372    377       4.
FT   REPEAT      381    386       5.
SQ   SEQUENCE   479 AA;  51600 MW;  990ED8376725DF61 CRC64;

  Query Match             3.6%;  Score 165;  DB 1;  Length 479;
  Best Local Similarity   23.1%;  Pred. No. 0.2;
  Matches  123;  Conservative   63;  Mismatches  203;  Indels  144;  Gaps   27;

Qy        370 LLVAGVTGLISGILEASKQAMFESVANRLQGKILEWEKQNGGQNYFDKGYDSRYAAYLAN 429

Db         40 LTVNGKTSY--SIDQAAAQITRENVS---------W----NGTNVFGKS---------AN 75

Qy        430 -NLKFLSELNKELEAERVIAITQQRWDNNIGELAGITKL-GERIKSGKAYADAFEDGKKV 487

Db         76 LTFKFLQSVSSIPSGD-----------------TGFVKFNAEQIEQAKLSLQSWSD---- 114

Qy        488 EAGSNITLDAKTGIIDISNSNGKKTQALHFTSPLLTAGTESRERLTNGKYSYINKLKFGR 547

Db        115 --VANLTFTEVTGNKSANITFGNYTRD---ASGNLDYGTQA--------YAYY------- 154

Qy        548 VKNWQVTDGEASSKLDFSKVIQRVAETEGTDEIG-LIVNAKAGNDDIFVGQGKMNIDGGD 606

Db        155 PGNYQ---GAGSSWYNYN---QSNIRNPGSEEYGRQTFTHEIGHALGLAHPGEYNAGEGD 208

Qy        607 -GHDRVFYSKD--------------------GGFGNITVDGTSATEAGSYTVNRKVARGD 645

Db        209 PSYNDAVYAEDSYQFSIMSYWGENETGADYNGHYGGAPMIDDIAAIQRLYGANMTTRTGD 268

Qy        646 IYHEVVKRQETKVGKRTETIQYRDYELRKVGYGYQSTDNLK----SVEEVIGSQFNDVFK 701

Db        269 SVY----------GFNSNT--DRDF--------YTATDSSKALIFSVWDAGGTDTFD-FS 307

Qy        702 GSKFNDIFHSGEG--DDL--LDG------GAGDDRLFGGKGNDRLSGDEGDDLLDGGSGD 751

Db        308 GYSNNQRINLNEGSFSDVGGLKGNVSIAHGVTIENAIGGSGNDILVGNSADNILQGGAGN 367

Qy        752 DVLNGGAGNDVYIFRKGDGNDTLYDGTGNDKLAFADANISDIMIERTKEGIIVKRND--- 808

Db        368 DVLYGGAGADT--LYGGAGRDTFVYGSGQDSTVAAYDWIADFQKGIDKIDLSAFRNEGQL 425

Qy        809 --------HSGSINIPRWYITSNLQNYQSNKTDH-KIEQLIGKDGSYITSDQI 852

Db        426 SFVQDQFTGKGQEVMLQWDAANSITNLWLHEAGHSSVDFLVRIVGQAAQSDII 478

RESULT 34
PRTG_ERWCH
```

```
ID   PRTG_ERWCH     STANDARD;      PRT;   475 AA.
AC   Q07162;
DT   01-FEB-1995 (Rel. 31, Created)
DT   01-FEB-1995 (Rel. 31, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Secreted protease G precursor (EC 3.4.24.-) (ProG).
GN   PRTG.
OS   Erwinia chrysanthemi.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Pectobacterium.
OX   NCBI_TaxID=556;
RN   [1]
RP   SEQUENCE FROM N.A., AND SEQUENCE OF N-TERMINUS.
RC   STRAIN=B374;
RX   MEDLINE=93234866; PubMed=1299839;
RA   Ghigo J.-M., Wandersman C.;
RT   "A fourth metalloprotease gene in Erwinia chrysanthemi.";
RL   Res. Microbiol. 143:857-867(1992).
CC   -!- COFACTOR: Binds 1 zinc ion (By similarity).
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- DOMAIN: THE GLY-RICH REPEATS MAY BE IMPORTANT IN THE EXTRACELLULAR
CC       SECRETION OF THIS METALLOPROTEASE.
CC   -!- SIMILARITY: Belongs to peptidase family M10B.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; X71365; CAA50501.1; -.
DR   PIR; S48132; S48132.
DR   HSSP; P23694; 1SAT.
DR   MEROPS; M10.055; -.
DR   InterPro; IPR001343; Hemlysn_Ca_bind.
DR   InterPro; IPR006026; NZn_MTpeptdse.
DR   InterPro; IPR006025; Zn_MTpeptdse.
DR   Pfam; PF00353; hemolysinCabind; 2.
DR   PRINTS; PR00313; CABNDNGRPT.
DR   SMART; SM00235; ZnMc; 1.
DR   PROSITE; PS00142; ZINC_PROTEASE; 1.
DR   PROSITE; PS00330; HEMOLYSIN_CALCIUM; 1.
KW   Hydrolase; Metalloprotease; Zinc; Zymogen; Calcium; Repeat.
FT   PROPEP        1     14
FT   CHAIN        15    475       SECRETED PROTEASE G.
FT   METAL       186    186       ZINC (CATALYTIC) (BY SIMILARITY).
FT   ACT_SITE    187    187       BY SIMILARITY.
FT   METAL       190    190       ZINC (CATALYTIC) (BY SIMILARITY).
FT   METAL       226    226       ZINC (CATALYTIC) (BY SIMILARITY).
FT   DOMAIN      341    382       4 X 6 AA REPEATS, GLY-RICH.
FT   REPEAT      341    346       1.
FT   REPEAT      359    364       2.
FT   REPEAT      370    375       3.
FT   REPEAT      377    382       4.
SQ   SEQUENCE   475 AA;  51577 MW;  2AD73D9A617F80B7 CRC64;

  Query Match             3.5%;  Score 164.5;  DB 1;  Length 475;
  Best Local Similarity   20.5%;  Pred. No. 0.21;
  Matches   99;  Conservative   58;  Mismatches  157;  Indels  169;  Gaps   21;

Qy        484 GKKVEAGSNITLDAKTGIIDI-------SNSNGKKTQALHFTSPLLTAGTESRERLTNGK 536
              |||  :  |: :    ||: |:          :  ||         :|       ::|:::| |
Db          5 GKKTDL-SSASSGGYTGVADVYQLWHYHARGNGNIGDKPSYTL------EQARDQITRGN 57

Qy        537 YSYINKLKFGRVKNWQTVDGEASSKLDFSKVIQRVAETEGTDEIGLIVNAKAGNDDIFVG 596
              ::   :   ||:           |: |  :|   :|  ||::: |     |    |     :
Db         58 ITWNGEKVFGK-----------SAALTYS-FLQSVADSDMPDNFKGFVKFNAAQ----IQ 101

Qy        597 QGKMNIDGGDGHDRVFYS--KDG-----GFGN--ITVDGTSATEA---GSYTVNRKVA-- 642
              | |: :          | :|  |||       |||  :| ||  :    :   | |  | | |
Db        102 QTKLALQSWADVANVTFSEAKDGERATIQFGNYTLTPDGNTDNNSQAFGFYPGNWKWAGS 161

Qy        643 ------------------RGDIYHEVVKRQETKVGKRTETIQYRDYELRKVGYGYQSTD 683
                                |  : ||          :|         ||:      ||:
Db        162 AWFNYNQADNQRPDINEFGRNTLTHE--------IGHTLGLYHPGDYDASDGNPGYKDVT 213

Qy        684 NLKSVEEV-IGSQFNDVFKGSKF--------NDI---------------------FHSGE 713
               .:  :   | | :|: :  |  |         :||                     |||
Db        214 YAEDTRQFSIMSYWNEGYTGGDFHGYHAAAPHDIAAIQKLYGANMSTRTGDTVYGFHSNS 273

Qy        714 GDD-------------LLDGGAGD---------------------------------- 724
              | |             : | | |
Db        274 GRDFYTATDSKTPLIFSVWDAGGNDTDFDSGYSANQRISLISGTFSDVGGLKAMVSIAAG 333

Qy        725 ---DRLFGGKGNDRLSGDEGDDLLDGGSGDDVLNGGAGNDVYIFRKGDGNDTLYDGTGND 781
                 :    || |:| : |:   |: :|||:|:|||  |   |  |   |   | | |     |
Db        334 AVIENAIGGSGHDVIVGNLSDNRIDGGAGNDVLYGDGGAD--ILTGGAGKDIFVYAWEKD 391

Qy        782 KLAFADANISDIMIERTKEGIIVKRNDH--------SGSIN--IPRWYITSNLQNYQSNK 831
                |: |    |:|        :  :     :|            ||  |   :  |    : |::|:|
Db        392 SLSSAPDTITDFQRGEDRIDLSAFNKNHDLRFVDNFSGKGNEVVLNW----DSQSHQTNM 447

Qy        832 TDH 834
               |
Db        448 WLH 450


RESULT 35
GTF2_STRDO
ID   GTF2_STRDO     STANDARD;      PRT;  1592 AA.
AC   P27470;
DT   01-AUG-1992 (Rel. 23, Created)
DT   01-AUG-1992 (Rel. 23, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Glucosyltransferase-I precursor (EC 2.4.1.5) (GTF-I) (Dextransucrase)
DE   (Sucrose 6-glucosyltransferase).
OS   Streptococcus downei (Streptococcus sobrinus).
OC   Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC   Streptococcus.
OX   NCBI_TaxID=1317;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=6715 / Serotype G;
RX   MEDLINE=91123227; PubMed=1704006;
RA   Abo H., Matsumura T., Kodama T., Ohta H., Fukui K., Kato K.,
RA   Kagawa H.;
RT   "Peptide sequences for sucrose splitting and glucan binding within
RT   Streptococcus sobrinus glucosyltransferase (water-insoluble glucan
RT   synthetase).";
RL   J. Bacteriol. 173:989-996(1991).
CC   -!- FUNCTION: PRODUCTION OF EXTRACELLULAR GLUCANS, THAT ARE THOUGHT
CC       TO PLAY A KEY ROLE IN THE DEVELOPMENT OF THE DENTAL PLAQUE BECAUSE
CC       OF THEIR ABILITY TO ADHERE TO SMOOTH SURFACES AND MEDIATE THE
CC       AGGREGATION OF BACTERIAL CELLS AND FOOD DEBRIS.
CC   -!- CATALYTIC ACTIVITY: Sucrose + {(1,6)-alpha-D-glucosyl}(N) = D-
CC       fructose + {(1,6)-alpha-D-glucosyl}(N+1).
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- DISEASE: DENTAL CARIES.
CC   -!- MISCELLANEOUS: GTF-I SYNTHESIZES WATER-INSOLUBLE GLUCANS (ALPHA
CC       1,3-LINKED GLUCOSE AND SOME 1,6 LINKAGES), GTF-S SYNTHESIZES
CC       WATER-SOLUBLE GLUCANS (ALPHA 1,6-GLUCOSE). GTF-SI SYNTHESIZES BOTH
CC       FORMS OF GLUCANS.
CC   -!- SIMILARITY: TO OTHER GLUCOSYLTRANSFERASES AND SOME TO A GLUCAN-
CC       BINDING PROTEIN FROM S.MUTANS.
CC   -!- SIMILARITY: Contains 16 cell wall binding repeats.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
```

```
DR   EMBL; D90213; BAA14241.1; -.
DR   InterPro; IPR002479; CW_binding.
DR   InterPro; IPR003318; Glyco_hydro_70.
DR   Pfam; PF01473; CW_binding_1; 13.
DR   Pfam; PF02324; Glyco_hydro_70; 1.
KW   Transferase; Glycosyltransferase; Signal; Repeat; Dental caries.
FT   SIGNAL         1     38       POTENTIAL.
FT   CHAIN         39   1592       GLUCOSYLTRANSFERASE-I.
FT   DOMAIN        39   1044       CATALYTIC (APPROXIMATE).
FT   DOMAIN      1093   1592       GLUCAN-BINDING (APPROXIMATE).
FT   DOMAIN      1093   1592       6.5 X TANDEM REPEATS.
FT   REPEAT      1093   1142       1.
FT   REPEAT      1158   1207       2.
FT   REPEAT      1222   1272       3.
FT   REPEAT      1287   1337       4.
FT   REPEAT      1402   1451       5.
FT   REPEAT      1514   1563       6.
FT   REPEAT      1577   1592       7 (INCOMPLETE).
SQ   SEQUENCE   1592 AA;  176167 MW;  BC0A66D079351ECF CRC64;

  Query Match              3.5%;  Score 163.5;  DB 1;  Length 1592;
  Best Local Similarity   21.0%;  Pred. No. 1.2;
  Matches  209;  Conservative  130;  Mismatches  384;  Indels  271;  Gaps   58;

Qy         16 NSTKSGL-----KNLYLAIPKD-----YDPQKGGTLNDFIK--AADELGIARLAEEPNHT 63
              :::|:||       :| ||     |        :||  ||  ::   |||: | |:|     :
Db        752 DASKAGLVKRTDENGYLYFLNDDLKGVANPQVSGFLQVWVPVGAADDQDI-RVAA----S 806

Qy         64 ETAKKSVDTVNQFLSLTQTGIAISATKLEKFLQKHSTNKLAKGLDSVENIDRKLGKASNV 123
              :||     :::|  ::                    :  : :|  : ::    |  : :||
Db        807 DTASTDGKSLHQDAAM-------------------DSRVMFEGFSNFQSFATKEEEYTNV 847

Qy        124 L--STLSSFLGTALAGIEL-------------DSLIKKGDAAPD----ALAKASIDLINE 164
              :  : :  |:   :   |:                ||:|: | |  |        ::||:
Db        848 VIANNVDKFVSWGITDFEMAPQYVSSTDGQFQDSVIQNGYAFTDRYDLGMSKAN------ 901

Qy        165 IIGNLSQSTQTIEAFSSQLAKLGSTISQAKGFSNIGNKL--QNLNFSKTNLGLEIITGLL 222
              |   |   : |:|                   |||   : : :   |    |     |::|
Db        902 KYGTADQLVKAIKAL------------HAKGLKVMADWVPDQMYTFPKQ----EVVT--- 942

Qy        223 SGISAGFALADKNASTGKKVAAGFELSNQVIGNVTKAISSYVLAQRVAAGL-STTGAVAA 281
                          :  || :  || :::: :         ||:         |:    |  |
Db        943 ---------VTRTDKFGKPI-AGSQINHSLYVTDTKSSGDDYQAKYGGAFLDELKEKYPE 992

Qy        282 LITSSIM---LAISPLAFMN--AADKFNHANALDEFAKQFRKFGYDGDHLLAEY--QRGV 334
              | |   |     || |    :   :|   || :| |                |    |::|::    :
Db        993 LFTKKQMSTGQAIDPSVKIKQWSAKYFNGSNIL----------GRGADYVLSDQVSNKYF 1042

Qy        335 GTIEASLTTISTALG-AVSAGVSAAAVG-----SAVGTPIALLVAGVTGLISGILEASKQ 388
                :|   |: ||  | :|:       |          || |   :             | | ||
Db       1043 NVASDTLFLPSSLLGKVVESGIRYDGKGYIYNSSATGDQVK---------ASFITEAGNL 1093

Qy        389 AMFESVANRLQGKILEWEKQNGGQNYF--DKGYDSRYAAY--LANNLKFLSELNKELEAE 444
              |          | ::     :   | |||  : |    |     |        |  :  :    |  | |
Db       1094 YYFGK-----DGYMVTGAQTINGANYFFLENGTALRNTIYTDAQGNSHYYANDGKRYENE 1148

Qy        445 RVIAITQQRW----DNNIGELAGITKLGERIKSGKAYADAFEDGKKVEAGSNITLDAKTG 500
                          |   |:      |:|  :   ::    |  |   :||  :      :|  | |
Db       1149 NGYQQFGNDWRYFKDGNMA--VGLTTVDGNVQ----YFD--KDGVQAKDKIIVTRDGKVR 1200

Qy        501 IIDISNSNGKKTQALHFTSPLLTAGTESRERLTNGKYSYINKLKFGRVKNWQVTDGEASS 560
                |  |  |           |:   :     |          |  : |: |            || |  :
Db       1201 YFDQHNGNA-------VTNTFIADKT--------GHWYYLGK-----------DGVAVT 1233

Qy        561 KLDFSKVIQRVAETEGTDEIGLIVNAKAGNDDIFVG-QGKM---NIDGGD-GHDRVFYSK 615
                           |:| |   ::     |   :      |      :||:    ::| ||    |      |
Db       1234 G----------AQTVGKQKLYFEANGGQVKGDFVTSDEGKLYFYDVDSGDMWTDTFIEDK 1283

Qy        616 DGGFGNITVDGTSATEAGS-------YTVNRKVARGDIYHEV---VKRQETKVGKR--TE 663
               | :  :  || : | | :              :   |  :  :|||          ::  : | |::     :
Db       1284 AGNWFYLGKDGAAVTGAQTIRGQKLYFKANGQQVKGDIVKGTDGKIRYYDAKSGEQVFNK 1343

Qy        664 TIQYRDYELRKVGYGYQSTDNLKSVEEVIGSQFNDVFKGSKFNDI---------------- 708
```

```
              |::   | :     :|      : |    ||   |  |   |  |          :|  ::
Db       1344 TVKAADGKTYVIGNDGVAVD--PSV--VKGQTFKDASGALRFYNLKGQLVTGSGWYETAN 1399

Qy        709 -----FHSGE---GDDLLDGGAGDDRLFGGKGNDRLSGD--EGDD----LLDGGSGDDVL 754
                   ||:    |:   ::|       |:   |     ::  |       |  |         |   |||
Db       1400 HDWVYIQSGKALTGEQTING----QHLYFKKDGHQVKGQLVTGTDGKVRYYDANSGDQAF 1455

Qy        755 NGG--AGNDVYIFRKGDGNDTLYDGTGNDK-LAFADANISDIMIERTKEGIIVKRNDHSG 811
                   |         |    |||          ||  |      |   |     |||      :  ||  :|      :|
Db       1456 NKSVTVNGKTYYF----GNDGTAQTAGNPKGQTFKDG--SDIRF-YSMEGQLV-----TG 1503

Qy        812 SINIPRWYITSNLQN---YQSNKTDHKIEQLIGKDGSYITSDQIDKILQDKKDGTVITSQ 868
              |        ||  || |     |     |          |  :|        |     :   |        |   |||
Db       1504 S----GWY--SNAQGQWLYVKNGKVLTGLQTVGSQRVYFDENGI-----QAKGKAVRTSD 1552

Qy        869 ELKKLADENKSQKLSASDIASSLNKLVGSMALFG 902
                 :  |||       | | |  :     ::  |       ||
Db       1553 GKIRYFDEN-----SGSMITNQWKEVNGRYYYFG 1581
```

RESULT 36
YMJB_CAEEL

```
ID   YMJB_CAEEL     STANDARD;      PRT;   918 AA.
AC   P34487;
DT   01-FEB-1994 (Rel. 28, Created)
DT   01-FEB-1995 (Rel. 31, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Hypothetical protein F59B2.12 in chromosome III.
GN   F59B2.12.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Bristol N2;
RX   MEDLINE=94150718; PubMed=7906398;
RA   Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
RA   Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
RA   Craxton M., Dear S., Du Z., Durbin R., Favello A., Fraser A.,
RA   Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M.,
RA   Johnston L., Jones M., Kershaw J., Kirsten J., Laisster N.,
RA   Latreille P., Lightning J., Lloyd C., Mortimore B., O'Callaghan M.,
RA   Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Shownkeen R.,
RA   Sims M., Smaldon N., Smith A., Smith M., Sonnhammer E., Staden R.,
RA   Sulston J., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K.,
RA   Waterston R., Watson A., Weinstock L., Wilkinson-Sproat J.,
RA   Wohldman P.;
RT   "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT   elegans.";
RL   Nature 368:32-38(1994).
CC   -----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----------------------------------------------------------------------
DR   EMBL; Z11505; CAA77581.1; -.
DR   PIR; G88545; G88545.
DR   WormPep; F59B2.12; CE01024.
KW   Hypothetical protein.
SQ   SEQUENCE   918 AA;  96560 MW;  E464FD86B14945DE CRC64;

  Query Match              3.5%;  Score 163;  DB 1;  Length 918;
  Best Local Similarity   18.5%;  Pred. No. 0.59;
  Matches  171;  Conservative  133;  Mismatches  358;  Indels  262;  Gaps   38;

Qy         83 GIAISATKLEKFLQKHSTNKLAKGLDSVENIDRKLGKASNVLSTLSSFLGTALAGIELDS 142
               | |:      |     ||::         |  :   :   | :||   |               :  |  :  :
Db         54 GFAMPKLDASKAAMVHSSS-------SHKGHHQSSGSSSNTHS-------LTVVGADGKN 99
```

```
Qy        143 LIKKGDAAPDALAKASIDLINE-----------IIGNLSQSTQTIEAFSSQLAKLGSTIS 191
              : :  :        :: :|  ||               |    |  :  : :| |  | |   | :
Db        100 ITENSEKKDGYNKESKVDEANENTKIKSADGSVIETGKSHNKSSDDASSYGLEKSSKTYA 159

Qy        192 QAKG--FSNIGNKLQNLNFSKTNLGLEIITGLLSGISAGFALADKNASTGKKVAAG-FEL 248
                   |  |:  ||: |                   |  || :  |     :: | | |
Db        160 DKNGTMLSSNTNKINNQ-------------------SRSAALDEGNEFVNQQNADGTFLR 200

Qy        249 SNQVIGNVTKAISSYVLAQRVAAGLSTTGAVAALITSSIMLAISPLAFMNAADKFNHANA 308
              :|     |  : :|  || :       :     |     :       :  |  |  :|   |    :|
Db        201 NNTGHKNTDEHLSHNVLDENAQMSIGADGTSHNITNRKGSVGDSHNAASDAHSNFESLDA 260

Qy        309 L-DEFAKQFRKFGYDGDHLLAEYQRGVGTIE-ASLTTISTALG----------------- 349
                :: :: :  |           |:::  : :::  |  |::|  : |
Db        261 QGNKKSQNYSKKAASASGSNADFESNLESLKNADGTSMSNSTGNFNNTSYDKATAEEVMS 320

Qy        350 -------------AVSAGVSAAAVGSAVGTPIALLVAGVTGLISGILEASKQAMFESVAN 396
                           |  || ::: :  ||  |       |  :  |    |  :  |            |  :|
Db        321 KKNVNADGTSSMEASHAGSNSSKINSASGQSSDLSMVGPNGIKS-----------HSTSN 369

Qy        397 RLQGKILEWEKQNGGQNYFDKGYDSRYAAYLANNLKFLSELNKELEAERVIAITQQRWDN 456
              :     |:    |:  |                                      :|::|
Db        370 KTDNYALDEANQSAG-----------------------------------SISEQ---- 389

Qy        457 NIGELAGITKLGERIKSGKAYADAFEDGKKVEAGSNI---TLDAKTGIIDISNSNGKKTQ 513
                      | | |:|          : |  |:|  |: :|        |||:        :|:|: |
Db        390 -------IGKNGQR----SLNESSIESGRKAESRNNTAADTLDSVDANGTVSSSHSKSAS 438

Qy        514 AL-----HFTSPLLTAGTESRERLTNGKY--SYINKL--KFGR------VKNWQVTDGEA 558
                      |  :  | |  :    :  |      || ||   :||         :||    |   :
Db        439 GTSLDENHNKTHALQASVDEHGNMKNHSIDGSYRNKKTGEFGNSEMSASIKNADGTMSQV 498

Qy        559 SSKLDFSKVIQRVAETEGTDEIGLIVNAKAGNDDIFVGQGKMNIDGGDGHDRVFYSKDGG 618
              : | | ::        |    ::   |  | :  :|   |  :  |        | :          |||
Db        499 NVKNDTNR------NTYEAEKSALEKNHEKNSDGTFKDESK-------GSNSRVNRTDGG 545

Qy        619 FGNITVDGTSATEAGSYTVNRKVARGDIYHEVVKRQETKVGKRTETIQYRDYELRKVGYG 678
               |: | |: :   |  :   :  |  :  |  : ::                   | |   :   :
Db        546 -SNLAV-GSYSVGKGGVSSNETIASSNAFNT------------------SDAESNQFDHL 585

Qy        679 YQSTDNLKSVEEVIGSQFNDVFKGSKFNDIFHSGEGDDLLDGGAGDDRLFGGKGN--DRL 736
              :| |  |         |::        | ||     :     |         ||        ||| |:
Db        586 HQKTAN--------GTEITHA-KDSK--QVAASANAKSSLDTSMS---AVDAKGNKVDKT 631

Qy        737 SGDEGDDLLDGGSGDDVLNGGAGNDVYIFRKGDGNDTLYDGTGN--------------D 781
              |     |         | :: :  | | :  |  :::: : : |                    :
Db        632 SSQAAD-------SHDAISASSDVDAKIVKHADRSESISNDSSNQTASEHNDSSKQSEHE 684

Qy        782 KLAFADANISDIMIERTKEGII-----VKRN-----DHSG-SINIPRWYITSNLQNYQSN 830
              |  || : ||:       |:   :          |::|    | :| |:  :   :    |   |
Db        685 KRQNADGSFSDVSSNSAKKNAVKDATDVRQNAFANVDAAGNSVSEVKNSLVENHSNDSDA 744

Qy        831 KTDHKIEQLIGKDG-------------SYITSDQIDKILQDKK----DGTVITSQELKKL 873
              :| ||        ||                   |::|  |: :|    ||      ||:|: |  :    :
Db        745 SSDSKIHSK-SVDGTELTDAKHSNVSTSHLTEGQLAQI--QKKLAMADGSVVMSNDNSHI 801

Qy        874 ADENKSQKLSASDIASSLNKLVGS 897
                  |     : |   |  |:|  ||
Db        802 THEKSRADVDAGHKA-HLSKADGS 824


RESULT 37
LMA3_HUMAN
ID   LMA3_HUMAN     STANDARD;      PRT;   1713 AA.
AC   Q16787; Q13679; Q13680;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Laminin alpha-3 chain precursor (Epiligrin 170 kDa subunit) (E170)
DE   (Nicein alpha subunit).
GN   LAMA3.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Keratinocytes;
RX   MEDLINE=94357926; PubMed=8077230;
RA   Ryan M.C., Tizard R., Vandevanter D.R., Carter W.G.;
RT   "Cloning of the LamA3 gene encoding the alpha 3 chain of the adhesive
RT   ligand epiligrin. Expression in wound repair.";
RL   J. Biol. Chem. 269:22779-22787(1994).
RN   [2]
RP   SEQUENCE OF 1-1331 FROM N.A. (ISOFORMS A AND B).
RX   MEDLINE=96163880; PubMed=8586427;
RA   Vidal F., Baudoin C., Miquel C., Galliano M.-F., Christiano A.M.,
RA   Uitto J., Ortonne J.-P., Meneguzzi G.;
RT   "Cloning of the laminin alpha 3 chain gene (LAMA3) and identification
RT   of a homozygous deletion in a patient with Herlitz junctional
RT   epidermolysis bullosa.";
RL   Genomics 30:273-280(1995).
CC   -!- FUNCTION: Binding to cells via a high affinity receptor, laminin
CC       is thought to mediate the attachment, migration, and organization
CC       of cells into tissues during embryonic development by interacting
CC       with other extracellular matrix components.
CC   -!- FUNCTION: LAMININ-5 IS THOUGHT TO BE INVOLVED IN (1) CELL ADHESION
CC       VIA INTEGRIN ALPHA-3/BETA-1 IN FOCAL ADHESION AND INTEGRIN ALPHA-
CC       6/BETA-4 IN HEMIDESMOSOMES, (2) SIGNAL TRANSDUCTION VIA TYROSINE
CC       PHOSPHORYLATION OF PP125-FAK AND P80, (3) DIFFERENTIATION OF
CC       KERATINOCYTES.
CC   -!- SUBUNIT: Laminin is a complex glycoprotein, consisting of three
CC       different polypeptide chains (alpha, beta, gamma), which are bound
CC       to each other by disulfide bonds into a cross-shaped molecule
CC       Comprising one long and three short arms with globules at each
CC       end. The alpha-3 chain is a subunit of laminin-5
CC       (epiligrin/kalinin/nicein), and possibly also a component of
CC       laminin-6 (K-laminin) and laminin-7 (KS-laminin).
CC   -!- SUBCELLULAR LOCATION: EXTRACELLULAR; FOUND IN THE BASEMENT
CC       MEMBRANES (MAJOR COMPONENT).
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=2;
CC       Name=A;
CC         IsoId=Q16787-1; Sequence=Displayed;
CC       Name=B;
CC         IsoId=Q16787-2; Sequence=VSP_003037;
CC         Note=Sequence incomplete;
CC   -!- TISSUE SPECIFICITY: SKIN; RESPIRATORY, URINARY, AND DIGESTIVE
CC       EPITHELIA AND IN OTHER SPECIALIZED TISSUES WITH PROMINENT
CC       SECRETORY OR PROTECTIVE FUNCTIONS. EPITHELIAL BASEMENT MEMBRANE,
CC       AND EPITHELIAL CELL TONGUE THAT MIGRATES INTO A WOUND BED. A
CC       DIFFERENTIAL AND FOCAL EXPRESSION OF THE ALPHA-3 CHAIN IS OBSERVED
CC       IN THE CNS.
CC   -!- INDUCTION: LAMININ-5 IS UP-REGULATED IN WOUND SITES OF HUMAN SKIN.
CC   -!- DOMAIN: THE ALPHA-HELICAL DOMAINS I AND II ARE THOUGHT TO INTERACT
CC       WITH OTHER LAMININ CHAINS TO FORM A COILED COIL STRUCTURE.
CC   -!- DOMAIN: DOMAIN G IS GLOBULAR.
CC   -!- DISEASE: DEFECTS IN LAMA3 ARE A CAUSE OF JUNCTIONAL EPIDERMOLYSIS
CC       BULLOSA (JEB) GRAVIS, ALSO KNOWN AS HERLITZ TYPE JEB. IT IS A
CC       BLISTERING DISORDER IN SKIN THAT IS CHARACTERIZED BY A SEPARATION
CC       OF BASAL CELLS FROM THE BASEMENT MEMBRANE DUE TO A DECREASED
CC       NUMBER OF HEMIDESMOSOMES. LAMININ-5 IS MISSING FROM THE BASEMENT
CC       MEMBRANE OF PATIENTS WITH THE GRAVIS FORM OF EPIDERMOLYSIS
CC       BULLOSA.
CC   -!- SIMILARITY: Contains 3 laminin EGF-like domains.
CC   -!- SIMILARITY: Contains 5 laminin G-like domains.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute.  There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
```

```
DR   EMBL; L34155; AAA59483.1; -.
DR   EMBL; X85107; CAA59428.1; -.
DR   EMBL; X85108; CAA59429.1; -.
DR   PIR; A55347; A55347.
DR   HSSP; P02468; 1TLE.
DR   Genew; HGNC:6483; LAMA3.
DR   MIM; 600805; -.
DR   MIM; 226700; -.
DR   GO; GO:0005604; C:basement membrane; TAS.
DR   GO; GO:0008544; P:epidermal differentiation; TAS.
DR   InterPro; IPR006209; EGF_like.
DR   InterPro; IPR002049; Laminin_EGF.
DR   InterPro; IPR001791; Laminin_G.
DR   Pfam; PF00053; laminin_EGF; 2.
DR   Pfam; PF00054; laminin_G; 2.
DR   SMART; SM00180; EGF_Lam; 2.
DR   SMART; SM00282; LamG; 5.
DR   PROSITE; PS00022; EGF_1; 1.
DR   PROSITE; PS01186; EGF_2; 1.
DR   PROSITE; PS01248; LAMININ_TYPE_EGF; 2.
DR   PROSITE; PS50025; LAM_G_DOMAIN; 5.
KW   Glycoprotein; Basement membrane; Extracellular matrix; Coiled coil;
KW   Laminin EGF-like domain; Cell adhesion; Repeat; Signal;
KW   Alternative splicing; Epidermolysis bullosa.
FT   SIGNAL        1     20       POTENTIAL.
FT   CHAIN        21   1713       LAMININ ALPHA-3 CHAIN.
FT   DOMAIN       46    201       DOMAIN III A.
FT   DOMAIN       67    113       LAMININ EGF-LIKE 1.
FT   DOMAIN      114    166       LAMININ EGF-LIKE 2.
FT   DOMAIN      167    185       LAMININ EGF-LIKE 3 (INCOMPLETE).
FT   DOMAIN      186    769       DOMAIN II AND I.
FT   DOMAIN      770    971       LAMININ G-LIKE 1.
FT   DOMAIN      978   1140       LAMININ G-LIKE 2.
FT   DOMAIN     1147   1307       LAMININ G-LIKE 3.
FT   DOMAIN     1366   1530       LAMININ G-LIKE 4.
FT   DOMAIN     1537   1710       LAMININ G-LIKE 5.
FT   DOMAIN      231    327       COILED COIL (POTENTIAL).
FT   DOMAIN      396    548       COILED COIL (POTENTIAL).
FT   DOMAIN      594    621       COILED COIL (POTENTIAL).
FT   DOMAIN      702    765       COILED COIL (POTENTIAL).
FT   DOMAIN     1686   1713       COILED COIL (POTENTIAL).
FT   DISULFID     67     76       BY SIMILARITY.
FT   DISULFID     69     83       BY SIMILARITY.
FT   DISULFID     86     95       BY SIMILARITY.
FT   DISULFID     98    111       BY SIMILARITY.
FT   DISULFID    114    126       BY SIMILARITY.
FT   DISULFID    116    135       BY SIMILARITY.
FT   DISULFID    137    146       BY SIMILARITY.
FT   DISULFID    149    164       BY SIMILARITY.
FT   DISULFID    202    202       INTERCHAIN (PROBABLE).
FT   DISULFID    205    205       INTERCHAIN (PROBABLE).
FT   SITE        658    660       CELL ATTACHMENT SITE (POTENTIAL).
FT   CARBOHYD    542    542       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    645    645       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    745    745       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    882    882       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    964    964       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1108   1108       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1131   1131       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1325   1325       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1477   1477       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1667   1667       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   VARSPLIC      1     45       MGWLWIFGAALGQCLGYSSQQQRVPFLQPPGQSQLQASYVE
FT                                FRPS -> KVSSYGGYLTYQAKSFALPGDMVLLEKKPDVQL
FT                                TGQHMSIIYEETNTPRPDRLHHGRVHVVEGNFRHASSRAPV
FT                                SREELMTVLSRLADVRIQGLYFTETQRLTLSEVGLEEASDT
FT                                GSGRIALAVEICACPPAYAGDSC (in isoform B).
FT                                /FTId=VSP_003037.
FT   CONFLICT      5      5       W -> R (IN REF. 2).
FT   CONFLICT    123    125       ATG -> GMC (IN REF. 2).
FT   CONFLICT    481    481       M -> K (IN REF. 2).
FT   CONFLICT    754    754       R -> L (IN REF. 2).
FT   CONFLICT    969    969       E -> Q (IN REF. 2).
FT   CONFLICT   1052   1052       D -> A (IN REF. 2).
FT   CONFLICT   1184   1184       G -> A (IN REF. 2).
SQ   SEQUENCE   1713 AA;  189304 MW;  45EA9BE1017B60D3 CRC64;

  Query Match             3.5%;  Score 163;  DB 1;  Length 1713;
  Best Local Similarity   20.0%;  Pred. No. 1.4;
  Matches  205;  Conservative  128;  Mismatches  333;  Indels  360;  Gaps   53;

Qy         30 PKDYDPQKG---------GTLNDFIKAADELGI-----------ARLAEEPNHTETAKKS 69
Db        190 PKDSSPAEECDDCDSCVMTLLNDLATMGEQLRLVKSQLQGLSASAGLLEQMRHMETQAK- 248

Qy         70 VDTVNQFLSLTQTGIAISATKLE-------------KFLQKHS----------TNKLAKG 106
Db        249 -DLRNQLLNY-RSAISNHGSKIEGLERELTDLNQEFETLQEKAQVNSRKAQTLNNNVNRA 306

Qy        107 LDSVENIDRKLGKA-SNVLSTLSSFLGTALAGIELDSLIKKGDAAPD------------- 152
Db        307 TQSAKELDVKIKNVIRNVHILLKQISGTDGEGNNVPS----GDFSREWAEAQRMMRELRN 362

Qy        153 -----ALAKASID------LINEI--------------IGNLSQSTQTIEAFSSQL-AKL 186
Db        363 RNFGKHLREAEADKRESQLLLNRIRTWQKTHQGENNGLANSIRDSLNEYEAKLSDLRARL 422

Qy        187 GSTISQAKGFSNIG-----------------NKLQNLNFSK-----------TNLGLEII 218
Db        423 QEAAAQAKQANGLNQENERALGAIQRQVKEINSLQS-DFTKYLTTADSSLLQTNIALQ-- 479

Qy        219 TGLLSGISAGFALADKNASTGKKVAAGFELSNQVIGNVTKAISSYVLAQRVAAGLSTTGA 278
Db        480 ------------LMEKSQKEYEKLAASLNEARQELSDKVRELSR-------SAGKT---- 516

Qy        279 VAALITSSIMLAISPLAFMNAADKFNHANALDEFAKQFR--KFGYDGDHLLAEYQRGVGT 336
Db        517 ----------------SLVEEAEK--HARSLQELAKQLEEIKRNASGDEL---------- 548

Qy        337 IEASLTTISTALGAVSAGVSAAAVGSAVGTPIALLVAGVTGLISGILEASKQAMFESVAN 396
Db        549 --------------VRCAVDAATA------------------YENILNAIKAA--EDAAN 574

Qy        397 RLQGKILEWEKQNGGQNYFDKGYDSRYAAYLANNLKFLSE---------------LNKEL 441
Db        575 RAAS-----ASESALQTVIKEDLPRKAKTLSSNSDKLLNEAKMTQKKLKQEVSPALNNLQ 629

Qy        442 EAERVIAITQQRWDNNIGELA----GITK--LGERIKSGKAYADAFEDGKKVEAGSNIT- 494
Db        630 QTLNIVTVQKEVIDTNLTTLRDGLHGIQRGDIDAMISSAKS---------MVRKANDITD 680

Qy        495 --LDA----KTGIIDISNSNGKKTQALHFTSPLLTAGTESRERLTNGKYSYINKL----- 543
Db        681 EVLDGLNPIQTDVERIKDTYG-RTQNEDF-KKALTDADNSVNKLT-------NKLPDLWR 731

Qy        544 KFGRVKNWQVTDGEASSKLD-FSKVIQRVAETEGTDEIGLIVNAKAG------ND----- 591
Db        732 KIESINQQLLPLGNISDNMDRIRELIQQARDAASKVAVPMRFNGKSGVEVRLPNDLEDLK 791

Qy        592 -----DIFVGQGKMNIDGGDGHDRVFY--SKDGG---FGNITVDG--TSATEAGSYTVNR 639
Db        792 GYTSLSLFLQRPNSRENGGTENMFVMYLGNKDASRDYIGMAVVDGQLTCVYNLGDREAEL 851

Qy        640 KVARGDIYHEVVKRQETKVGKRTETIQYRDYELRKVGYGYQSTDN---LKSVEEVIGSQF 696
Db        852 QV------DQILTKSETKEAVMDRVKFQRIYQFARLNYTKGATSSKPETPGVYDMDGRNS 905

Qy        697 NDVFKGSKFNDIFHSGEGDDLLDGGAGDDRLFGGKGNDRLSGD------EGDDLLDGGSG 750
Db        906 NTLLNLDPENVVFYVG-------GYPPDFKL-----PSRLSFPPYKGCIELDDL-----N 948

Qy        751 DDVLNGGAGNDVYIF---------------RKGDGNDTLYDGTGNDKL---AFADANIS 791
Db        949 ENVLS------LYNFKKTFNLNTTEVEPCRRRKEESDKNYFEGTGYARVPTQPHAPIPTF 1002

Qy        792 DIMIERTKE-GII--VKRNDHSGSINIP------RWYITSNLQNYQ------SNKTDHKI 836
```

```
             |: | : |:: : | |:||    |: : | |  :    :| || |
Db      1003 GQTIQTTVDRGLLFFAENGDRFISLNIEDGKLMVRYKLNSELPKERGVGDAINNGRDHSI 1062

Qy       837 EQLIGK 842
             : |||
Db      1063 QIKIGK 1068

RESULT 38
YDBA_ECOLI
ID   YDBA_ECOLI     STANDARD;      PRT;  2003 AA.
AC   P33666; P76087; P76088; P76856; P76857; P76859;
DT   01-FEB-1994 (Rel. 28, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Hypothetical protein ydbA.
GN   YDBA OR B1401/B1405.
OS   Escherichia coli.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Escherichia.
OX   NCBI_TaxID=562;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K12 / MG1655;
RX   MEDLINE=97426617; PubMed=9278503;
RA   Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA   Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA   Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA   Mau B., Shao Y.;
RT   "The complete genome sequence of Escherichia coli K-12.";
RL   Science 277:1453-1474(1997).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K12;
RX   MEDLINE=97251357; PubMed=9097039;
RA   Aiba H., Baba T., Fujita K., Hayashi K., Inada T., Isono K., Itoh T.,
RA   Kasai H., Kashimoto K., Kimura S., Kitakawa M., Kitagawa M.,
RA   Makino K., Miki T., Mizobuchi K., Mori H., Mori T., Motomura K.,
RA   Nakade S., Nakamura Y., Nashimoto H., Nishio Y., Oshima T., Saito N.,
RA   Sampei G., Seki Y., Sivasundaram S., Tagami H., Takeda J.,
RA   Takemoto K., Takeuchi Y., Wada C., Yamamoto Y., Horiuchi T.;
RT   "A 570-kb DNA sequence of the Escherichia coli K-12 genome
RT   corresponding to the 28.0-40.1 min region on the linkage map.";
RL   DNA Res. 3:363-377(1996).
RN   [3]
RP   SEQUENCE OF 464-2003 FROM N.A.
RC   STRAIN=K12;
RX   MEDLINE=92190338; PubMed=1665988;
RA   Moszer I., Glaser P., Danchin A.;
RT   "Multiple IS insertion sequences near the replication terminus in
RT   Escherichia coli K-12.";
RL   Biochimie 73:1361-1374(1991).
CC   -!- SIMILARITY: TO S.TYPHIMURIUM ORF NEAR CYSG (AC P25928).
CC   -!- CAUTION: THIS IS A CONCEPTUAL TRANSLATION; THE GENE CODING FOR
CC       THIS PROTEIN IS INTERRUPTED BY A HYBRID IS2D/IS30C ELEMENT
CC       BETWEEN AMINO ACIDS 839 AND 840.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; AE000237; AAC74483.1; ALT_SEQ.
DR   EMBL; AE000237; AAC74487.1; ALT_SEQ.
DR   EMBL; D90778; BAA15009.1; ALT_SEQ.
DR   EMBL; D90778; BAA18880.1; ALT_SEQ.
DR   EMBL; D90779; BAA18881.1; ALT_SEQ.
DR   EMBL; X62680; -; NOT_ANNOTATED_CDS.
DR   EcoGene; EG11307; ydbA.
KW   Hypothetical protein; Complete proteome.
FT   CONFLICT    489    489       I -> V (IN REF. 2).
FT   CONFLICT    495    495       I -> V (IN REF. 2).
SQ   SEQUENCE   2003 AA;  205949 MW;  B83A12C8B53220EE CRC64;

  Query Match             3.5%;  Score 162.5;  DB 1;  Length 2003;
  Best Local Similarity   20.4%;  Pred. No. 1.8;
  Matches  124;  Conservative   75;  Mismatches  197;  Indels  213;  Gaps   28;

Qy       430 NLKFLSELNKELEAERVIAITQQRWDNNIGELAGITKLGERIKSGKAYADAFEDGKKVEA 489
             || : | |  : | | |   :|  |          || : :|         |   |
Db       185 NLWQIDEANNTVALEGVSADGATKWQYNHN--------GELVITG--------DNATVNN 228

Qy       490 GSNITLDAK--TGIIDISNSNGKKTQ-----------ALHFTSPLLTAGTESRERLT--- 533
                |:| |  || :|: :||| |               :  |     |   :   :|
Db       229 NGKTTVDGKDSTG-TEINGNNGKVIQDGDLDVSGGGHGIDITGDSATVDNKGTMTVTDPE 287

Qy       534 ------NGKYSYINKLKFGRVKN----WQVTDGEASSKLDFSKVIQRVAETEGTDEIGLI 583
                   :|  : :|      : |     |:   :|:: :   :       :| |  |
Db       288 SMGIQIDGDKAIVNNEGESTITNGGTGTQINGDDATANNNGKTTV------DGKDSTGTE 341

Qy       584 VNAKAGNDDIFVGQGKMNIDGG-----------------------------DGHDRVF 612
             :|  ||:   :   | ::: ||                              || |
Db       342 IN---GNNGKVIQDGDLDVSGGGHGIDITGDSATVDNKGTMTVTDPESIGIQVDGDQAVV 398

Qy       613 YSK------DGGFG-NITVDGTSATEAGSYTVNRKVARGDIYHEVVKRQETKV-GKRTET 664
                ::      :|| |  |  |   :|   |   ||:  |  :|     |:: |   :
Db       399 NNEGESAITNGGTGTQINGDDATANNNGKTTVDGKDSTG-----------TEIAGNNGKV 447

Qy       665 IQYRDYELRKVGYGYQSTDNLKSVE----------EVIGSQFNDVFKGSKFNDIFHSGEG 714
             || | ::  |:|  | : :|:          | || |                    :|
Db       448 IQDGDLDVSGGGHGIDITGDSATVDNKGTMTVTDPESIGIQI---------------DG 491

Qy       715 DDLLDGGAGDDRLFGG------KGNDRLSGDEGDDLLDG--GSGDDVLNGGAGNDVYIFR 766
             |  :    |: :  |        ||| : : |   :||    :|  :    |||   :
Db       492 DQAIVNNEGESTITNGGTGTQINGNDATANNSGKTTVDGKDSTGTKI----AGNIGIV-- 545

Qy       767 KGDGNDTLYDGTGNDKLAFADANISDIMIERTKEGIIVKRNDHSGSINIPRWYITSNLQN 826
               ||: |:  |         | ||  |    |  |:|        :|||  :
Db       546 NLDGSLTVTGG------AHGVENIGDNGTVNNKGDIVV---SDTGSIGV----------- 585

Qy       827 YQSNKTDHKIEQLIGKDGSYITSDQIDKILQDKKDGTVITSQELK--------------- 871
                           || :|: : |:  |  : :   | ||: |
Db       586 ------------LINGEGATV-SNTGDVNVSNEATGFSITTNSGKVSLAGSMQVGDFSTG 632

Qy       872 -KLADENKSQKLSASDIASSLNKLVGSMA----LFGTANSV------------SSNALQ 913
                |  | |  |:| |:    |:||   |    : | ||:|              : || :
Db       633 VDLNGNNNSVTLAAKDL-----KVVGQKATGINVSGDANTVNITGNVLVDKDKTADNAAE 687

Qy       914 PITQPTQGI 922
                 |: ||
Db       688 YFFDPSVGI 696

RESULT 39
TOXB_CLODI
ID   TOXB_CLODI     STANDARD;      PRT;  2366 AA.
AC   P18177;
DT   01-NOV-1990 (Rel. 16, Created)
DT   01-NOV-1990 (Rel. 16, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Toxin B.
GN   TOXB OR TCDB.
OS   Clostridium difficile.
OC   Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC   Clostridium.
OX   NCBI_TaxID=1496;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=VPI 10463;
RX   MEDLINE=90326540; PubMed=2374729;
RA   Barroso L.A., Wang S.Z., Phelps C.J., Johnson J.L., Wilkins T.D.;
RT   "Nucleotide sequence of Clostridium difficile toxin B gene.";
RL   Nucleic Acids Res. 18:4004-4004(1990).
```

```
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=VPI 10463;
RA   von Eichel-Streiber C.;
RL   Submitted (JAN-1997) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE OF 1271-2366 FROM N.A.
RC   STRAIN=VPI 10463;
RX   MEDLINE=92293124; PubMed=1603068;
RA   Eichel-Streiber C., Laufenberg-Feldmann R., Sartingen S., Schulze J.,
RA   Sauerborn M.;
RT   "Comparative sequence analysis of the Clostridium difficile toxins A
RT   and B.";
RL   Mol. Gen. Genet. 233:260-268(1992).
CC   -!- DISEASE: CLOSTRIDIUM DIFFICILE PRODUCES TWO TOXINS, AN ENTEROTOXIN
CC       CALLED A AND CYTOTOXIN B.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; X53138; CAA37298.1; -.
DR   EMBL; X92982; CAA63562.1; -.
DR   EMBL; X60984; CAA43299.1; -.
DR   PIR; A27636; A27636.
DR   PIR; S10317; S10317.
DR   InterPro; IPR002479; CW_binding.
DR   Pfam; PF01473; CW_binding_1; 17.
DR   Pfam; PF04488; Gly_transf_sug; 1.
KW   Toxin.
SQ   SEQUENCE   2366 AA;  269709 MW;  E1024BD8B8A56ADF CRC64;

  Query Match             3.5%;  Score 161;  DB 1;  Length 2366;
  Best Local Similarity   18.4%;  Pred. No. 2.6;
  Matches  223;  Conservative  164;  Mismatches  375;  Indels  448;  Gaps   55;

Qy          3 NINVIKSNIQAGLNSTKSGLKNLYLAIPKDYDPQKGGTLNDFIKAADELGIARLAEEPNH 62
Db        667 DVDSLSTEIEAAIDLAKEDIS------PKSIE-------------INLLGCNMFSYSINV 707

Qy         63 TET-----AKKSVDTVNQFL-SLTQTGIAISATKLE--------KFLQKHS---TNK--- 102
Db        708 EETYPGKLLLKVKDKISELMPSISQDSIIVSANQYEVRINSEGRRELLDHSGEWINKEES 767

Qy        103 LAKGLDSVENI-----DRKLGKASNVLSTLSSFL-----GTALAGIELDSLIKKGDAAPD 152
Db        768 IIKDISSKEYISFNPKENKITVKSKNLPELSTLLQEIRNNSNSSDIELEEKVMLTECEIN 827

Qy        153 ALAKASIDLINEIIGNLSQSTQTIEAFSSQLAKLGSTISQAKGFSNIGNKLQNLNF---- 208
Db        828 VISNIDTQIVEERIEEAKNLTSDSINYIKDEFKLIESISDALCDLKQQNELEDSHFISFE 887

Qy        209 --SKTNLGLEI-------------------------ITGLLSGISAGFALADKNASTGK 240
Db        888 DISETDEGFSIRFINKETGESIFVETEKTIFSEYANHITEEISKIK-GTIFDTVNGKLVK 946

Qy        241 KV------------AAGF-------ELSNQVIGNVTKAISSYVLAQRVAAGLST---TGA 278
Db        947 KVNLDTTHEVNTLNAAFFIQSLIEYNSSKESLSNLSVAMKVQVYAQLFSTGLNTITDAAK 1006

Qy        279 VAALITSSIMLAISPLAFMNAA--------DKFNHANALDEFAKQFRKFGYDGDHLL-AE 329
Db       1007 VVELVSTALDETIDLLPTLSEGLPIIATIIDGVSLGAAIKELSE-------TSDPLLRQE 1059

Qy        330 YQRGVGTIEASLTTISTALGAVSAGVSAAAVGSAVGTPIALLVAGVTGLISGIL----EA 385
Db       1060 IEAKIGIMAVNLTTATTAIITSSLGIASGF--SILLVPLAGISAGIPSLVNNELVLRDKA 1117

Qy        386 SKQAMF---------ESVANRLQGKILE-----------------------WEKQNGGQ 412
Db       1118 TKVVDYFKHVSLVETEGVFTLLDDKIMMPQDDLVISEIDFNNNSIVLGKCEIWRMEGGSG 1177

Qy        413 NYFDKGYDSRYAA----YLANNLKFLS---------ELNKEL-----EAERVIAITQQRW 454
Db       1178 HTVTDDIDHFFSAPSITYREPHLSIYDVLEVQKEELDLSKDLMVLPNAPNRVFA-----W 1232

Qy        455 DNNIGELAGI-------TKLGERIK---SGKAY-------ADA--------FEDGKKVEA 489
Db       1233 ET--GWTPGLRSLENDGTKLLDRIRDNYEGEFYWRYFAFIADALITTLKPRYED-----T 1285

Qy        490 GSNITLDAKTGIIDISNSNGKKTQALHFTSPLLTAGTE-SRERL------TNGKYSYINK 542
Db       1286 NIRINLDSNT---------------RSFIVPIIT--TEYIREKLSYSFYGSGGTYA-LSL 1327

Qy        543 LKFGRVKNWQVTDGEA---------------SSKLDFSKVIQRVAETEGTDEIGLIVNAK 587
Db       1328 SQYNMGINIELSESDVWIIDVDNVVRDVTIESDKIKKGDLIEGILSTLSIEENKIILNSH 1387

Qy        588 AGNDDIFVGQGKMNIDGGDGHDRVFYS------------------------------- 614
Db       1388 EIN---FSGE----VNGSNGFVSLTFSILEGINAIIEVDLLSKSYKLLISGELKILMLNS 1440

Qy        615 -----------------------------KDGGFGN-ITVDGTSATEAGSYTVNRKVARG 644
Db       1441 NHIQQKIDYIGFNSELQKNIPYSFVDSEGKENGFINGSTKEGLFVSELPDVVLISKVYMD 1500

Qy        645 DIYHEVVKRQETKVGKRTETIQYRDYELRKVGYGYQSTDNLKSVEEV--------IGSQF 696
Db       1501 D--------------------------SKPSFGYYS-NNLKDVKVITKDNVNILTGYYL 1532

Qy        697 NDVFKGS-----------KFNDIFHSGEGDDLLDGGAGDDRLFGGKGNDRLSGD------ 739
Db       1533 KDDIKISLSLTLQDEKTIKLNSV-HLDE-----SGVAEILKFMNRKGNTNTSDSLMSFLE 1586

Qy        740 --------------------EGDDLLDG----GSGDDVLNGGAGNDVYI--FRKGDGNDT 773
Db       1587 SMNIKSIFVNFLQSNIKFILDANFIISGTTSIGQFEFICDENDNIQPYFIKFNTLETNYT 1646

Qy        774 LYDGTGNDKLAFADANISD----------------IMIERTKEGIIVKRNDHSGSINIPR 817
Db       1647 LYVGNRQNMIVEPNYDLDDSGDISSTVINFSQKYLYGIDSCVNKVVISPNIYTDEINITP 1706

Qy        818 WYITSNL---------QNYQSNKTDHKIEQLI---------GKDGSYITSDQIDKILQDK-- 859
Db       1707 VYETNNTYPEVIVLDANYINEKINVNINDLSIRYVWSNDGNDFILMSTSEENKVSQVKIR 1766

Qy        860 -----KDGTV 864
Db       1767 FVNVFKDKTL 1776

RESULT 40
PM21_CHLPN
ID   PM21_CHLPN     STANDARD;      PRT;  1609 AA.
AC   Q9Z6U5; Q9RB58;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Probable outer membrane protein pmp21 precursor (Polymorphic membrane
DE   protein 21).
GN   PMP21 OR CPN0963 OR CP0897.
OS   Chlamydia pneumoniae (Chlamydophila pneumoniae).
OC   Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydophila.
OX   NCBI_TaxID=83558;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CWL029;
RX   MEDLINE=99206606; PubMed=10192388;
RA   Kalman S., Mitchell W., Marathe R., Lammel C., Fan J., Hyman R.W.,
RA   Olinger L., Grimwood J., Davis R.W., Stephens R.S.;
RT   "Comparative genomes of Chlamydia pneumoniae and C. trachomatis.";
RL   Nat. Genet. 21:385-389(1999).
```

```
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=AR39;
RX   MEDLINE=20150255; PubMed=10684935;
RA   Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
RA   White O., Hickey E.K., Peterson J., Utterback T., Berry K., Bass S.,
RA   Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,
RA   Gwinn M., Nelson W., DeBoy R., Kolonay J., McClarty G., Salzberg S.L.,
RA   Eisen J., Fraser C.M.;
RT   "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia
RT   pneumoniae AR39.";
RL   Nucleic Acids Res. 28:1397-1406(2000).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=J138;
RX   MEDLINE=20330349; PubMed=10871362;
RA   Shirai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,
RA   Shiba T., Ishii K., Hattori M., Kuhara S., Nakazawa T.;
RT   "Comparison of whole genome sequences of Chlamydia pneumoniae J138
RT   from Japan and CWL029 from USA.";
RL   Nucleic Acids Res. 28:2311-2314(2000).
CC   -!- SUBCELLULAR LOCATION: CELL WALL SURFACE (ELEMENTARY BODIES)
CC       (POTENTIAL).
CC   -!- SIMILARITY: BELONGS TO THE PMP OUTER MEMBRANE PROTEIN FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; AE001676; AAD19099.1; -.
DR   EMBL; AE002248; AAF38684.1; -.
DR   EMBL; AP002548; BAA99171.1; -.
DR   PIR; A86611; A86611.
DR   PIR; H72013; H72013.
DR   PHCI-2DPAGE; Q9RB58; -.
DR   TIGR; CP0897; -.
DR   InterPro; IPR006315; Autotransport.
DR   InterPro; IPR005546; Autotransporter.
DR   InterPro; IPR003368; Chlamydia_PMP.
DR   Pfam; PF03797; Autotransporter; 1.
DR   Pfam; PF02415; DUF145; 1.
DR   TIGRFAMs; TIGR01414; autotrans_barl; 1.
DR   TIGRFAMs; TIGR01376; POMP_repeat; 13.
KW   Outer membrane; Signal; Multigene family; Complete proteome.
FT   SIGNAL        1     30       POTENTIAL.
FT   CHAIN        31   1609       PROBABLE OUTER MEMBRANE PROTEIN PMP21.
FT   CONFLICT    420    420       I -> M (IN REF. 3).
SQ   SEQUENCE   1609 AA;  170865 MW;  2604C3E9FC4024CB CRC64;

  Query Match             3.5%;  Score 160.5;  DB 1;  Length 1609;
  Best Local Similarity   21.0%;  Pred. No. 1.7;
  Matches  244;  Conservative  129;  Mismatches  446;  Indels  343;  Gaps   59;

Qy         17 STKSGLKNLYLAIPKDYDPQKGGTLNDFIKAADELGIARLAEEPNH--TETAKKSVD--- 71
Db         72 SQKESEKVLYTQVPLTQ-----GSSGESLDLAD----ANFLEHFQHLFEETTVFGIDQKL 122

Qy         72 -----TVNQFLSLTQ---TGIAISATKLEKFLQKHSTNKLAKGLDSVENIDRKLGKASNV 123
Db        123 VWSDLDTRNFSQPTQEPDTSNAVS----EKISSDTKENR--KDLET-EDPSKKSG-LKEV 174

Qy        124 LSTLSSFLGTALAGIELDSLIKKGDAAPDALAKASIDLINEIIGNLSQSTQTIEA--FSS 181
Db        175 SSDLPKSPETAVAAISEDLEISENISARDPLQGLAFFYKNTSSQSISEKDSSFQGIIFS- 233

Qy        182 QLAKLGSTISQAKGFSNI----------------GNKLQNLNFSK------------- 210
Db        234 -----GSGANSGLGFENLKAPKSGAAVYSDRDIVFENLVKGLSFISCESLEDGSAAGVNI 288

Qy        211 --TNLGLEIITGLLSGISAGFALADKNASTGKKV--AAGFELSNQVIGNVTKAISS--YV 264
Db        289 VVTHCGDVTLTDCATGLDLEALRLVKDFSRGGAVFTARNHEVQNNLAGGILSVVGNKGAI 348

Qy        265 LAQRVAAGLSTTGAVA-----------ALITSSIMLAISPLAFMNAAD---------K 302
Db        349 VVEKNSAEKSNGGAFACGSFVYSNNENTALWKEN--QALSGGAISSASDIDIQGNCSAIE 406

Qy        303 FNHANALDEFAKQFRKFGYDGDHLLAEY----QRGVGTIEASLTTISTALGAVSAGV--- 355
Db        407 FSGNQSLIALGEHIGLTDFVGGGALAAQGTLTLRNNAVVQCVKNTSKTHGGAILAGTVDL 466

Qy        356 -----------SAAAVGSAVGTPIALLVAGVTGLISGILEASKQAMFESVANRLQGKIL 403
Db        467 NETISEVAFKQNTAALTGGALSANDKVIIANNFG----------EILFEQNEVRNHGGAI 516

Qy        404 --------EWEKQNGGQNYFDKGYDSRYAAYLANNLKFLSELNKELEAERVIAITQQRW- 454
Db        517 YCGCRSNPKLEQKDSGENINIIG-NSGAITFLKNKASVLEVMTQAEDYAGGGAL----WG 571

Qy        455 -----DNNIGELAGITKLGERIKSGKAYADAFEDGKKVEAGSNITLDAKTGIIDISNSN- 508
Db        572 HNVLLDSNSGNIQFIGNIG---------GSTFWIGEYVGGGAILSTDRVT----ISNNSG 618

Qy        509 -----GKKTQAL--HFTSPLLTAGTESRERLTNGKYSYINKLKFGRVKNWQVTDGEA-SS 560
Db        619 DVVFKGNKGQCLAQKYVAPQETAPVESDASSTNKDEKSLNACSHGDHYPPKTVEEEVPPS 678

Qy        561 KLDFSKVI------------QRVAETEGTDEIGLIVNAKAGNDDIFVGQGKMNIDGGD- 606
Db        679 LLEEHPVVSSTDIRGGGAILAQHIFITDNTGNLRFSGNLGGGEESSTVGD--LAIVGGGA 736

Qy        607 ----------GHDRVFYSKDGGFGNITVDGTSATEAGSYTVNRKVARGDI--YH--EVVK 652
Db        737 LLSTNEVNVCSNQNVVFS-----DNVTSNG---CDSGGAILAKKV---DISANHSVEFVS 785

Qy        653 RQETKVG----KRTETIQYRD------YELRKVGYGYQSTDNLKSVEEVIGSQFNDVFKG 702
Db        786 NGSGKFGGAVCALNESVNITDNGSAVSFSKNRTRLGGAGVAAPQGSVTICGNQGNIAFK- 844

Qy        703 SKFNDIFHSGEGDDLLDGGA-----------GDDRLF-----GGKGN----DRLSGDEG 741
Db        845 ----ENFVFGSENQRSGGGAIIANSSVNIQDNAGDILFVSNSTGSYGGAIFVGSLVASEG 900

Qy        742 DD----LLDGGSGD----------------DVLNGGA--------------------- 758
Db        901 SNPRTLTITGNSGDILFAKNSTQTAASLSEKDSFGGGAIYTQNLKIVKNAGNVSFYGNRA 960

Qy        759 --GNDVYIFRKGD------GNDTLYDGT------------GNDK-----LAFADANI-- 790
Db        961 PSGAGVQIADGGTVCLEAFGGDILFEGNINFDGSFNAIHLCGNDSKIVELSAVQDKNIIF 1020

Qy        791 SDIMI--ERTKEGIIVKRNDHSGSINIPRWYITSNLQNYQSNKTDHKIEQLIGKDGSYIT 848
Db       1021 QDAITYEENTIRGLPDK---DVSPLSAPSLIFNSKPQD---DSAQH-------HEGTIRF 1067

Qy        849 SDQIDKILQDK--KDGTVITSQE----LKKLADENKSQ-KLSASDIASSLNKLVGSMALF 901
Db       1068 SRGVSKIPQIAAIQEGTLALSQNAELWLAGLKQETGSSIVLSAGSILRIFDSQVDSSAPL 1127

Qy        902 GTANS----VSSNALQPITQPT 919
Db       1128 PTENKEETLVSAGVQINMSSPT 1149

Search completed: February 17, 2004, 10:13:48
Job time : 31 secs
```

```
OM protein - protein search, using sw model

Run on:         February 17, 2004, 10:09:49 ; Search time 27 Seconds
                                          (without alignments)
                                          3301.791 Million cell updates/sec

Title:          US-10-069-799-5
Perfect score:  4647
Sequence:       1 MSNINVIKSNIQAGLNSTKS..........SSNALQPITQPTQGILAPSV 927

Scoring table:  BLOSUM62
                Gapop 10.0 , Gapext 0.5

Searched:       283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters:      283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       PIR_76:*
                1:  pir1:*
                2:  pir2:*
                3:  pir3:*
                4:  pir4:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.
```

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2325	50.0	953	1	B30169	leukotoxin A - Pas
2	2236.5	48.1	956	1	B33389	toxin II - Actinob
3	2236	48.1	955	1	A35254	leukotoxin A - Pas
4	2205.5	47.5	956	1	A43834	toxin II - Actinob
5	1955.5	42.1	1052	1	B49219	toxin III - Actino
6	1924	41.4	1049	1	S51784	toxin III - Actino
7	1847.5	39.8	1024	2	S10056	hemolysin A - Esch
8	1845	39.7	1023	1	LEECA	hemolysin A - Esch
9	1816	39.1	1055	1	A37205	leukotoxin A - Act
10	1763.5	37.9	1022	1	I39643	RTX-toxin I - Acti
11	1748	37.6	998	2	T00227	hemolysin A toxin
12	1729	37.2	998	2	I41078	hemolysin - Escher
13	805.5	17.3	1705	2	S51672	adenylate cyclase
14	802	17.3	1706	1	OYBRC	cyclolysin - Borde
15	437	9.4	1829	2	E81086	iron-regulated pro
16	435	9.4	1829	2	S35027	cytotoxin RTX homo
17	421	9.1	1302	2	C81182	iron-regulated pro
18	415	8.9	1636	2	B82736	hemolysin-type cal
19	411	8.8	1208	2	C82779	hemolysin-type cal
20	408.5	8.8	1115	2	A47058	Fe-regulated RTX c
21	391	8.4	2064	2	G82562	bacteriocin XF2407
22	383	8.2	208	2	S34238	leukotoxin A - Pas
23	366	7.9	697	2	F81856	probable RTX-famil
24	344	7.4	4936	2	AH2515	hypothetical prote
25	330.5	7.1	1944	2	AH3098	rhizobiocin/RTX to
26	330.5	7.1	1990	2	A96188	probable phosphoes
27	329	7.1	1296	2	C82521	hemolysin-type cal
28	305.5	6.6	993	2	AE1905	outer membrane sec
29	301.5	6.5	1072	2	G95851	probable hemolysin
30	288	6.2	1417	2	AG2137	hypothetical prote
31	287.5	6.2	3083	2	AH2493	hypothetical prote
32	280	6.0	643	2	T03518	hypothetical prote
33	280	6.0	900	2	AC1852	hypothetical prote
34	276	5.9	650	2	G87572	calcium-binding pr
35	269.5	5.8	1112	2	H95964	probable outer mem
36	268	5.8	539	2	G95405	hypothetical prote
37	267.5	5.8	387	2	E95995	hypothetical calci
38	266.5	5.7	589	2	AD2263	hypothetical prote
39	266	5.7	2348	2	AD1841	hypothetical prote
40	263.5	5.7	387	2	E95933	probable calcium-b
41	262.5	5.6	1428	2	AC2224	hypothetical prote
42	262.5	5.6	1741	2	S74910	hemolysin - Synech
43	262.5	5.6	1839	2	S77626	mannuronan C-5-epi
44	262.5	5.6	3262	2	AH2137	hypothetical prote
45	261	5.6	1028	2	T03516	probable outer mem

ALIGNMENTS

```
RESULT 1
B30169
leukotoxin A - Pasteurella haemolytica (serotype 1)
N;Alternate names: lktA protein
C;Species: Pasteurella haemolytica
C;Date: 12-Oct-1989 #sequence_revision 15-Nov-1996 #text_change 18-Jun-1999
C;Accession: B30169; C32051; S29516
R;Highlander, S.K.; Chidambaram, M.; Engler, M.J.; Weinstock, G.M.
DNA 8, 15-28, 1989
A;Title: DNA sequence of the Pasteurella haemolytica leukotoxin gene cluster.
A;Reference number: A30169; MUID:89210283; PMID:2707120
A;Accession: B30169
A;Status: not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 1-953 <HIG>
R;Strathdee, C.A.; Lo, R.Y.C.
J. Bacteriol. 171, 916-928, 1989
A;Title: Cloning, nucleotide sequence, and characterization of genes encoding the secret
A;Reference number: A32051; MUID:89123172; PMID:2914876
A;Accession: C32051
A;Status: not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 947-953 <STR>
R;Lo, R.Y.C.; Strathdee, C.A.; Shewen, P.E.
Infect. Immun. 55, 1987-1996, 1987
A;Title: Nucleotide sequence of the leukotoxin genes of Pasteurella haemolytica A1.
A;Reference number: S29515; MUID:87306837; PMID:3040588
A;Accession: S29516
A;Molecule type: DNA
A;Residues: 1-741,'D',743-953 <LOR>
A;Cross-references: EMBL:M20730; NID:g150492; PIDN:AAA25529.1; PID:g150494
C;Comment: This organism causes bovine pneumonic pasteurellosis (shipping fever).
C;Genetics:
A;Gene: lktA
C;Function:
A;Description: lyses leukocytes
C;Superfamily: hemolysin A; hemolysin A homology
C;Keywords: calcium binding; cytolysis; exotoxin; hemolysis; lipoprotein; tandem repeat;
F;238-784/Domain: hemolysin A homology <HLYA>
F;716-807/Region: 9-residue repeats (G-G-X-G-[DN]-D-X-[LVIYF]-X)
F;554/Binding site: palmitate (Lys) (covalent) #status predicted

  Query Match             50.0%;  Score 2325;  DB 1;  Length 953;
  Best Local Similarity   49.9%;  Pred. No. 7.9e-112;
  Matches  464;  Conservative  175;  Mismatches  264;  Indels   26;  Gaps   13;

Qy          8 KSNIQAGLNSTKSGLKNLYLAIPKD--YDPQKGGTLNDFIKAADELGIARLAEEPNHTET 65
              :|  ||| :| |:| | : | ||::  || ::|  | | :|||:||||    || |:  |
Db         28 QSLTQAG-SSLKTGAKKIILYIPQNYQYDTEQGNGLQDLVKAAEELGIEVQREERNNIAT 86

Qy         66 AKKSVDTVNQFLSLTQTGIAISATKLEKFLQKHSTNKLAKGLDSVENIDRKLGKASNVLS 125
              |: |: |:   : ||: || :|| :::| |||    |  : | | |:| :   ||  |||
```

```
Db        87 AQTSLGTIQTAIGLTERGIVLSAPQIDKLLQK---TKAGQALGSAESIVQNANKAKTVLS 143

Qy       126 TLSSFLGTALAGIELDSLIKKGDAAPDALAKASIDLINEIIGNLSQSTQTIEAFSSQLAK 185
Db       144 GIQSILGSVLAGMDLDEAL-QNNSNQHALAKAGLELTNSLIENIANSVKTLDEFGEQISQ 202

Qy       186 LGSTISQAKGFSNIGNKLQNL-NFSKTNLGLEIITGLLSGISAGFALADKNASTGKKVAA 244
Db       203 FGSKLQNIKGLGTLGDKLKNIGGLDKAGLGLDVISGLLSGATAALVLADKNASTAKKVGA 262

Qy       245 GFELSNQVIGNVTKAISSYVLAQRVAAGLSTTGAVAALITSSIMLAISPLAFMNAADKFN 304
Db       263 GFELANQVVGNITKAVSSYILAQRVAAGLSSTGPVAALIASTVSLAISPLAFAGIADKFN 322

Qy       305 HANALDEFAKQFRKFGYDGDHLLAEYQRGVGTIEASLTTISTALGAVSAGVSAAAVGSAV 364
Db       323 HAKSLESYAERFKKLGYDGDNLLAEYQRGTGTIDASVTAINTALAAIAGGVSAAAAGSVI 382

Qy       365 GTPIALLVAGVTGLISGILEASKQAMFESVANRLQGKILEWEKQNGGQNYFDKGYDSRYA 424
Db       383 ASPIALLVSGITGVISTILQYSKQAMFEHVANKIHNKIVEWEKNNHGKNYFENGYDARYL 442

Qy       425 AYLANNLKFLSELNKELEAERVIAITQQRWDNNIGELAGITKLGERIKSGKAYADAFEDG 484
Db       443 ANLQDNMKFLLNLNKELQAERVIAITQQQWDNNIGDLAGISRLGEKVLSGKAYVDAFEEG 502

Qy       485 KKVEAGSNITLDAKTGIIDISNSNGKKTQALHFTSPLLTAGTESRERLTNGKYSYINKLK 544
Db       503 KHIKADKLVQLDSANGIIDVSNSGKAKTQHILFRTPLLTPGTEHRERVQTGKYEYITKLN 562

Qy       545 FGRVKNWQVTDGEASSKLDFSKVIQRV-------AETEGTDEIGLIVNAKAGNDDIFVGQ 597
Db       563 INRVDSWKITDGAASSTFDLTNVVQRIGIELDNAGNVTKTKETKIIAKLGEGDDNVFVGS 622

Qy       598 GKMNIDGGDGHDRVFYSKDGGFGNITVDGTSATEAGSYTVNRKVARGDIYHEVVKRQETK 657
Db       623 GTTEIDGGEGYDRVHYSR-GNYGALTIDATKETEQGSYTVNRFVETGKALHEVTSTHTAL 681

Qy       658 VGKRTETIQYRDYELRKVGYGYQSTDNLKSVEEVIGSQFNDVFKGSKFNDIFHSGEGDDL 717
Db       682 VGNREEKIEYR-HSNNQHHAGYYTKDTLKAVEEIIGTSHNDIFKGSKFNDAFNGGDGVDT 740

Qy       718 LDGGAGDDRLFGGKGNDRLSGDEGDDLLDGGSGDDVLNGGAGNDVYIFRKGDGNDTLYDG 777
Db       741 IYGNDGNDRLFGGKGDDILDGGNGDDFIDGGKGNDLLHGGKGDDIFVHRKGDGNDIITDS 800

Qy       778 TGNDKLAFADANISDIMIERTKEGIIVKRNDHSGSINIPRWY----ITSNLQNYQSNKTD 833
Db       801 DGNDKLSFSDSNLKDLTFEKVKHNLVI-TNSKKEKVTIQNWFREADFAKEVPNYKATK-D 858

Qy       834 HKIEQLIGKDGSYITSDQIDKILQDKKDGTVITSQELKKLADENKSQKLSASDIASSLNK 893
Db       859 EKIEEIIGQNGERITSKQVDDLI--AKGNGKITQDELSKVVDNYELLKHS-KNVTNSLDK 915

Qy       894 LVGSMALFGTANSVSSNALQPITQPTQGI 922
Db       916 LISSVSAFTSSNDSRNVLVAPTSMLDQSL 944
```

RESULT 2
B33389
toxin II - Actinobacillus pleuropneumoniae
N;Alternate names: cytolysin II; RTX-toxin II (ApxII)
C;Species: Actinobacillus pleuropneumoniae
C;Date: 09-Mar-1990 #sequence_revision 01-Nov-1996 #text_change 18-Jun-1999
C;Accession: B33389; S18853; B43599
R;Chang, Y.F.; Young, R.; Struck, D.K.
DNA 8, 635-647, 1989
A;Title: Cloning and characterization of a hemolysin gene from Actinobacillus (Haemophil
A;Reference number: A33389; MUID:90126233; PMID:2693022
A;Accession: B33389
A;Molecule type: DNA
A;Residues: 1-956 <CHA>
A;Cross-references: GB:M30602; NID:g141823; PIDN:AAA87232.1; PID:g141825
A;Experimental source: serotype 5
R;Smits, M.A.; Briaire, J.; Jansen, R.; Smith, H.E.; Kamp, E.M.; Gielkens, A.L.J.
submitted to the EMBL Data Library, July 1991.
A;Description: Cytolysins of Actinobacillus pleuropneumoniae serotype 9.
A;Reference number: S18852
A;Accession: S18853
A;Molecule type: DNA
A;Residues: 1-956 <SMI>
A;Cross-references: EMBL:X61111; NID:g38939; PIDN:CAA43423.1; PID:g38941
R;Smits, M.A.; Briaire, J.; Jansen, R.; Smith, H.E.; Kamp, E.M.; Gielkens, A.L.J.
Infect. Immun. 59, 4497-4504, 1991
A;Title: Cytolysins of Actinobacillus pleuropneumoniae serotype 9.
A;Reference number: A43599; MUID:92040145; PMID:1937809
A;Accession: B43599
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-27;948-956 <SM2>
A;Cross-references: GB:X61111; NID:g38939
C;Comment: This organism causes porcine pleuropneumonia.
C;Genetics:
A;Gene: apxIIA; appA; clyIIA
C;Function:
A;Description: attacks blood cell membranes and causes cell lysis
C;Superfamily: hemolysin A; hemolysin A homology
C;Keywords: calcium binding; cytolysis; exotoxin; hemolysis; lipoprotein; tandem repeat;
F;243-787/Domain: hemolysin A homology <HLYA>
F;719-801/Region: 9-residue repeats (G-G-X-G-[DN]-D-X-[LVIYF]-X)
F;557/Binding site: palmitate (Lys) (covalent) #status predicted

```
  Query Match             48.1%;  Score 2236.5;  DB 1;  Length 956;
  Best Local Similarity   47.7%;  Pred. No. 2.8e-107;
  Matches  453;  Conservative  188;  Mismatches  255;  Indels   53;  Gaps   15;

Qy         1 MSNINV--IKSNIQAGL---------------------NSTKSGLKNLYLAIPKDYDPQ 36
Db         1 MSKITLSSLKSSLQQGLKNGKNKLNQAGTTLKNGLTQTGHSLQNGAKKLILYIPQGYDSG 60

Qy        37 KGGTLNDFIKAADELGIARLAEEPNHTETAKKSVDTVNQFLSLTQTGIAISATKLEKFLQ 96
Db        61 QGNGVQDLVKAANDLGIEVWREERSNLDIAKTSFDTTQKILGFTDRGIVLFAPQLDNLLK 120

Qy        97 KHSTNKLAKGLDSVENIDRKLGKASNVLSTLSSFLGTALAGIELDSLIKKGDAAPDALAK 156
Db       121 KNP--KIGNTLGSASSISQNIGKANTVLGGIQSILGSVLSGVNLNELLQNKDPNQLELAK 178

Qy       157 ASIDLINEIIGNLSQSTQTIEAFSSQLAKLGSTISQAKGFSNIGNKLQNL-NFSKTNLGL 215
Db       179 AGLELTNELVGNIASSVQTVDAFAEQISKLGSHLQNVKGLGGLSNKLQNLPDLGKASLGL 238

Qy       216 EIITGLLSGISAGFALADKNASTGKKVAAGFELSNQVIGNVTKAISSYVLAQRVAAGLST 275
Db       239 DIISGLLSGASAGLILADKEASTEKKAAAGVEFANQIIGNVTKAVSSYILAQRVASGLSS 298

Qy       276 TGAVAALITSSIMLAISPLAFMNAADKFNHANALDEFAKQFRKFGYDGDHLLAEYQRGVG 335
Db       299 TGPVAALIASTVALAVSPLSFLNVADKFKQADLIKSYSERFQKLGYDGDRLLADFHRETG 358

Qy       336 TIEASLTTISTALGAVSAGVSAAAVGSAVGTPIALLVAGVTGLISGILEASKQAMFESVA 395
Db       359 TIDASVTTINTALAAISGGVGAASAGSLVGAPVALLVAGVTGLITTILEYSKQAMFEHVA 418

Qy       396 NRLQGKILEWEKQNGGQNYFDKGYDSRYAAYLANNLKFLSELNKELEAERVIAITQQRWD 455
Db       419 NKVHDRIVEWEKKH-NKNYFEQGYDSRHLADLQDNMKFLINLNKELQAERVVAITQQRWD 477

Qy       456 NNIGELAGITKLGERIKSGKAYADAFEDGKKVEAGSNITLDAKTGIIDISNSNGKKTQAL 515
Db       478 NQIGDLAAISRRTDKISSGKAYVDAFEEGQHQSYDSSVQLDNKNGIINISNTN-RKTQSV 536

Qy       516 HFTSPLLTAGTESRERLTNGKYSYINKLKFGRVKNWQVTDGEASSKLDFSKVIQRVA--- 572
Db       537 LFRTPLLTPGEENRERIQEGKNSYITKLHIQRVDSWTVTDGDASSSVDFTNVVQRIAVKF 596
```

```
Qy   573 -------ETEGTDEIGLIVNAKAGNDDIFVGQGKMNIDGGDGHDRVFYSKDGGFGNITVD 625
Db   597 DDAGNIIESKDTK---IIANLGAGNDNVFVGSSTTVIDGGDGHDRVHYSR-GEYGALVID 652

Qy   626 GTSATEAGSYTVNRKVARGDIYHEVVKRQETKVGKRTETIQYRDYELRKVGYGYQSTDNL 685
Db   653 ATAETEKGSYSVKRYVGDSKALHETIATHQTNVGNREEKIEYR-REDDRFHTGYTVTDSL 711

Qy   686 KSVEEVIGSQFNDVFKGSKFNDIFHSGEGDDLLDGGAGDDRLFGGKGNDRLSGDEGDDLL 745
Db   712 KSVEEIIGSQFNDIFKGSQFDDVFHGGNGVDTIDGNDGDDHLFGGAGDDVIDGGNGNNFL 771

Qy   746 DGGSGDDVLNGGAGNDVYIFRKGDGNDTLYDGTGNDKLAFADANISDIMIERTKEGIIVK 805
Db   772 VGGTGNDIISGGKDNDIYVHKTGDGNDSITDSGGQDKLAFSDVNLKDLTFKKVDSSLEI- 830

Qy   806 RNDHSGSINIPRWY----ITSNLQNYQSNKTDHKIEQLIGKDGSYITSDQIDKILQDKKD 861
Db   831 INQKGEKVRIGNWFLEDDLASTVANYKAT-NDRKIEEIIGKGGERITSEQVDKLI--KEG 887

Qy   862 GTVITSQELKKLADENKSQKLSASDIASSLNKLVGSMALFGTANSVSSN 910
Db   888 NNQISAEALSKVVNDYNTSK-DRQNVSNSLAKLISSVGSFTSSSDFRNN 935
```

RESULT 3
A35254
leukotoxin A - Pasteurella haemolytica (serotype T10)
N;Alternate names: lktA protein
C;Species: Pasteurella haemolytica
C;Date: 10-Aug-1990 #sequence_revision 15-Nov-1996 #text_change 18-Jun-1999
C;Accession: S37145; A35254; S34237; S34235
R;Lainson, A.F.; Aitchison, K.; Donachie, W.
submitted to the EMBL Data Library, September 1993
A;Description: DNA sequence of the leukotoxin A gene from P. haemolytica T10 serotype.
A;Reference number: S37145
A;Accession: S37145
A;Molecule type: DNA
A;Residues: 1-955 <LAI>
A;Cross-references: EMBL:Z26247; NID:g400424; PIDN:CAA81206.1; PID:g400425
R;Highlander, S.K.; Engler, M.J.; Weinstock, G.M.
J. Bacteriol. 172, 2343-2350, 1990
A;Title: Secretion and expression of the Pasteurella haemolytica leukotoxin.
A;Reference number: A35254; MUID:90236888; PMID:2185213
A;Accession: A35254
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 950-955 <HIG>
A;Cross-references: GB:M24197; GB:M34943; GB:M34944
R;Lainson, A.F.; Aitchison, K.D.; Donachie, W.
submitted to the EMBL Data Library, June 1993
A;Description: DNA sequence of the carboxy terminal end of leukotoxin A from the T3 sero
A;Reference number: S34235
A;Accession: S34237
A;Molecule type: DNA
A;Residues: 745-955 <LA2>
A;Cross-references: EMBL:Z22884; NID:g311828; PIDN:CAA80498.1; PID:g311829
A;Experimental source: serotype T3
A;Accession: S34235
A;Molecule type: DNA
A;Residues: 723-955 <LA3>
A;Cross-references: EMBL:Z22887; NID:g311824; PIDN:CAA80501.1; PID:g311825
A;Experimental source: serotype T10
C;Function:
A;Description: attacks cell membranes and causes cell lysis
C;Superfamily: hemolysin A; hemolysin A homology
C;Keywords: calcium binding; cytolysis; exotoxin; hemolysis; lipoprotein; tandem repeat;
F;240-786/Domain: hemolysin A homology <HLYA>
F;718-809/Region: 9-residue repeats (G-G-X-G-[DN]-D-X-[LVIYF]-X)
F;718-726/Region: repeat
F;727-735/Region: repeat
F;736-744/Region: repeat
F;745-753/Region: repeat
F;754-762/Region: repeat
F;763-771/Region: repeat
F;772-780/Region: repeat
F;781-789/Region: repeat
F;792-800/Region: repeat
F;801-809/Region: repeat
F;556/Binding site: palmitate (Lys) (covalent) #status predicted

```
  Query Match             48.1%;  Score 2236;  DB 1;  Length 955;
  Best Local Similarity   48.2%;  Pred. No. 3e-107;
  Matches  458;  Conservative  178;  Mismatches  271;  Indels   44;  Gaps   17;

Qy     5 NVIKSNIQA---GLN-----------STKSGLKNLYLAIPKD--YDPQKGGTLNDFIKAA 48
Db    12 NGIRSTLTATRGGLNRAGQSLTQAGQTLKNGAKKIILYIPKDYKYDSGSGNGLQDLVKAA 71

Qy    49 DELGIARLAEEPNHTETAKKSVDTVNQFLSLTQTGIAISATKLEKFLQKHSTNKLAKGLD 108
Db    72 EELGIEVQKEEGNDIAKAQTSLGTIQNVLGLTERGIVLSAPQLDKLLQK---NKVGQALG 128

Qy   109 SVENIDRKLGKASNVLSTLSSFLGTALAGIELDSLIKKGDAAPDALAKASIDLINEIIGN 168
Db   129 SSESTAQNFSQAKTVLSGVQGNSRTVLAGMDLDEAL-QNESDQLTLAKAGLELTNSLIEN 187

Qy   169 LSQSTQTIEAFSSQLAKLGSTISQAKGFSNIGNKLQNL-NFSKTNLGLEIITGLLSGISA 227
Db   188 IANSVQTLDAFSEQISQFGSKLQNVKGLGALGDKLKNIGGLDKAGLGLDVKSRLLSGATA 247

Qy   228 GFALADKNASTGKKVAAGFELSNQVIGNVTKAISSYVLAQRVAAGLSTTGAVAALITSSI 287
Db   248 ALVLADKDASTAKKVGAGFELANQVVGNITKAVSSYILAQRVAAGLSSTGPVAALIASTV 307

Qy   288 MLAISPLAFMNAADKFNHANALDEFAKQFRKFGYDGDHLLAEYQRGVGTIEASLTTISTA 347
Db   308 AVAISPLSFAGIADKFDRAKSLENYAERFKKLGYEGDSLLAEYQHGTGTIDASVTAINTA 367

Qy   348 LGAVSAGVSAAAVGSAVGTPIALLVAGVTGLISGILEASKQAMFESVANRLQGKILEWEK 407
Db   368 LAAIAGGVSAAAAGSVVASPIALLVSGITGVISTILQYSKQAMFEHVANKIHNKIVEWEK 427

Qy   408 QNGGQNYFDKGYDSRYAAYLANNLKFLSELNKELEAERVIAITQQRWDNNIGELAGITKL 467
Db   428 NNGGKNYFENGYDARYLANLQDNMKFLLNLNKELQAERVIAITQQQWDSNIGDLAGISRL 487

Qy   468 GERIKSGKAYADAFEDGKKVEAGSNITLDAKTGIIDISNSNGKKTQALHFTSPLLTAGTE 527
Db   488 GEKVLSGKAYVDAFEEGQHLKADKLVQLDSAKGIIDVTNTGEAKTQHILFRTPLLTPGTE 547

Qy   528 SRERLTNGKYSYINKLKFGRVKNWQVTDGEASSKLDFSKVIQRVA-------ETEGTDEI 580
Db   548 KRERVQTGKYEYITKLHINRVDSWQIKDGAASSTFDLTNVVQRIGVELDHAENVIKTKET 607

Qy   581 GLIVNAKAGNDDIFVGQGKMNIDGGDGHDRVFYSKDGGFGNITVDGTSATEAGSYTVNRK 640
Db   608 KIVATLGDGDDNVFVGSGTTEIDGGEGYDRVHYSR-GNYGALTIDATKETEQGSYTVNRF 666

Qy   641 VARGDIYHEVVKRQETKVGKRTETIQYRDYELRKVGYGYQSTDNLKSVEEVIGSQFNDVF 700
Db   667 VESGKALHEGTSTHTALVGNREEKIEYR-HSNNQHHAGYYTKDTLKAVEEIIGTSHNDIF 725

Qy   701 KGSKFNDIFHSGEGDDLLDGGAGDDRLFGGKGNDRLSGDEGDDLLDGGSGDDVLNGGAGN 760
Db   726 KGSKFNDAFNGGDGVDTIDGNDGNDRLFGGKGDDIIDGGNGDDFIDGGKGNDLLHGGKGD 785

Qy   761 DVYIFRKGDGNDTLYDGTGNDKLAFADANISDIMIERTKEGIIVKRNDHSGSINIPRWY- 819
Db   786 DIFVHRQGDGNDSITESEGNDKLSFSDSNLKDLTFEKVNHHLVI-TNTKQEKVTIQNWFR 844

Qy   820 ---ITSNLQNYQSNKTDHKIEQLIGKDGSYITSDQIDKILQDKKDGTVITSQELKKLADE 876
Db   845 EAEFAKTIQNYVATR-DDKIEEIIGQNGERITSKQVDELIE--KGNGKIAQSELTKVVDN 901
```

```
Qy        877 NKSQKLSASDIASSLNKLVGSMALFGTANSVSSNALQPITQPTQGILAPSV 927

Db        902 YQLLKYS-RDASNSLDKLISSASAFTSSND-SRNVL---ASPT-SMLDPSL 946

RESULT 4
A43834
toxin II - Actinobacillus suis
N;Alternate names: ashA protein; cytolysin II; RTX-toxin II
C;Species: Actinobacillus suis
C;Date: 31-Dec-1993 #sequence_revision 08-Nov-1996 #text_change 05-Dec-1998
C;Accession: A43834
R;Burrows, L.L.; Lo, R.Y.
Infect. Immun. 60, 2166-2173, 1992
A;Title: Molecular characterization of an RTX toxin determinant from Actinobacillus suis
A;Reference number: A43834; MUID:92267623; PMID:1587585
A;Accession: A43834
A;Molecule type: DNA
A;Residues: 1-956 <BUR>
A;Experimental source: isolate 3714
A;Note: sequence extracted from NCBI backbone (NCBIN:104212, NCBIP:104211)
C;Comment: This organism causes acute fatal septicemia in young pigs.
C;Function:
A;Description: attacks cell membranes and causes cell lysis
C;Superfamily: hemolysin A; hemolysin A homology
C;Keywords: calcium binding; cytolysis; exotoxin; hemolysis; lipoprotein; tandem repeat;
F;243-787/Domain: hemolysin A homology <HLYA>
F;719-801/Region: 9-residue repeats (G-G-X-G-[DN]-D-X-[LVIYF]-X)
F;557/Binding site: palmitate (Lys) (covalent) #status predicted

  Query Match             47.5%;  Score 2205.5;  DB 1;  Length 956;
  Best Local Similarity   47.5%;  Pred. No. 1.1e-105;
  Matches  450;  Conservative  187;  Mismatches  260;  Indels   51;  Gaps   15;

Qy          1 MSNINV--IKSNIQAGL---------------------NSTKSGLKNLYLAIPKDYDPQ 36

Db          1 MSKITLSSLKSSLQQGLKNGKNKLNQAGTTLKNGLTQTGHSLQNGAKKLILYIPQGYDSG 60

Qy         37 KGGTLNDFIKAADELGIARLAEEPNHTETAKKSVDTVNQFLSLTQTGIAISATKLEKFLQ 96

Db         61 QGNGVQDLVKAANDLGIEVWREERSNLDIAKTSFDTTQKILGFTDRGIVLFAPQLKNLLK 120

Qy         97 KHSTNKLAKGLDSVENIDRKLGKASNVLSTLSSFLGTALAGIELDSLIKKGDAAPDALAK 156

Db        121 KNP--KIGNTLGSASSISQNIGKANTVLGGIQSILGSVLSGVNLNELLQNKDPNQLELAK 178

Qy        157 ASIDLINEIIGNLSQSTQTIEAFSSQLAKLGSTISQAKGFSNIGNKLQNL-NFSKTNLGL 215

Db        179 AGLELTNELVGNIASSVQTVDAFAEQISKLGSHLQNVKGLGGLSNKLQNLPDLGKASLGL 238

Qy        216 EIITGLLSGISAGFALADKNASTGKKVAAGFELSNQVIGNVTKAISSYVLAQRVAAGLST 275

Db        239 DIISGLLSGASAGLILADKKASTEKKAAAGVEFANQIIGNVTKAVSSYILAQRVASCLSS 298

Qy        276 TGAVAALITSSIMLAISPLAFMNAADKFNHANALDEFAKQFRKFGYDGDHLLAEYQRGVG 335

Db        299 TGPVAALIASTVALAVSPLSFLNVADKFKQADLIKSYSERFQKLGYDGDRLLADFHRETG 358

Qy        336 TIEASLTTISTALGAVSAGVSAAAVGSAVGTPIALLVAGVTGLISGILEASKQAMFESVA 395

Db        359 TIDASVTTINTALAAISGGVGAASAGSLVGAPVALLVAGVTGLITTILEYSKQAMFEHVA 418

Qy        396 NRLQGKILEWEKQNGGQNYFDKGYDSRYAAYLANNLKFLSELNKELEAERVIAITQQRWD 455

Db        419 NKVHDRIVEWEKKH-NKNYFEQGYDSRHLADLQDNMKFLINLNKELQAERVVAITQQRWD 477

Qy        456 NNIGELAGITKLGERIKSGKAYADAFEDGKKVEAGSNITLDAKTGIIDISNSNGKKTQAL 515

Db        478 NQIGDLAAISRRTDKISSGKAYVDAFEEGQHQSYDSSVQLDNKNGIINISNTN-RKTQSV 536

Qy        516 HFTSPLLTAGTESRERLTNGKYSYINKLKFGRVKNWQVTDGEASSKLDFSKVIQRVAETE 575

Db        537 LFRTPLLTPGEENRERIQEGKNSTITKLHIQRVDSWTVTVGDASSSVDFTNVVQR--EAV 594

Qy        576 GTDEIGLIVNAK---------AGNDDIFVGQGKMNIDGGDGHDRVFYSKDGGFGNITVDG 626

Db        595 KFDDAGNIIESKDTKIIANLGAGNDNVFVGSSTTVIDGGDGHDRVHYSR-GEYGALVIDA 653

Qy        627 TSATEAGSYTVNRKVARGDIYHEVVKRQETKVGKRTETIQYRDYELRKVGYGYQSTDNLK 686

Db        654 TAETEKGSYSVKRYVGDSKALHETIATHQTNVGNREEKIEYR-REDDRFHTGYTVTDSLK 712

Qy        687 SVEEVIGSQFNDVFKGSKFNDIFHSGEGDDLLDGGAGDDRLFGGKGNDRLSGDEGDDLLD 746

Db        713 SVEEIIGSQFNDIFKGSQFDDVFHGGNGVDTIDGNDGDDHLFGGAGDDVIDGGNGNNFLV 772

Qy        747 GGSGDDVLNGGAGNDVYIFRKGDGNDTLYDGTGNDKLAFADANISDIMIERTKEGIIVKR 806

Db        773 GGTGNDIISGGKDNDIYVHKTGDGNDSITDSGGQDKLAFSDVNLKDLTFKKVDSSLEI-I 831

Qy        807 NDHSGSINIPRWY----ITSNLQNYQSNKTDHKIEQLIGKDGSYITSDQIDKILQDKKDG 862

Db        832 NQKGEKVRIGNWFLEDDLASTVANYKAT-NDRKIEEIIGKGGERITSEQVDKLI--KEGN 888

Qy        863 TVITSQELKKLADENKSQKLSASDIASSLNKLVGSMALFGTANSVSSN 910

Db        889 NQISAEALSKVVNDYNTSK-DRQNVSNSLAKLISSVGSFTSSSDFRNN 935

RESULT 5
B49219
toxin III - Actinobacillus pleuropneumoniae (serotype 8)
N;Alternate names: RTX-toxin III (ApxIIIA)
C;Species: Actinobacillus pleuropneumoniae
C;Date: 19-Dec-1993 #sequence_revision 01-Nov-1996 #text_change 18-Jun-1999
C;Accession: B49219; S48043; S29958
R;Jansen, R.; Briaire, J.; Kamp, E.M.; Gielkens, A.L.; Smits, M.A.
Infect. Immun. 61, 947-954, 1993
A;Title: Cloning and characterization of the Actinobacillus pleuropneumoniae-RTX-toxin I
A;Reference number: A49219; MUID:93162836; PMID:8432615
A;Accession: B49219
A;Molecule type: DNA
A;Residues: 1-1052 <JAN1>
A;Cross-references: EMBL:X68815; NID:g38956; PIDN:CAA48711.1; PID:g38958
A;Experimental source: strain 405, serotype 8
A;Note: sequence extracted from NCBI backbone (NCBIN:125168, NCBIP:125170)
R;Jansen, R.; Briaire, J.; van Geel, A.B.M.; Kamp, E.M.; Gielkens, A.L.J.; Smits, M.A.
Infect. Immun. 62, 4411-4418, 1994
A;Title: Genetic map of the Actinobacillus pleuropneumoniae RTX-toxin (Apx) operons: cha
A;Reference number: S48042; MUID:95012630; PMID:7927703
A;Accession: S48043
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-1052 <JAN2>
A;Cross-references: EMBL:X80055; NID:g558150; PIDN:CAA56358.1; PID:g558152
A;Experimental source: strain 405, serotype 8
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, July 1994
C;Comment: This orgnism causes porcine pleuropneumonia.
C;Genetics:
A;Gene: apxIIIA
C;Function:
A;Description: lyses lung macrophages
C;Superfamily: hemolysin A; hemolysin A homology
C;Keywords: calcium binding; cytolysis; exotoxin; lipoprotein; tandem repeat; thiolester
F;254-804/Domain: hemolysin A homology <HLYA>
F;736-862/Region: 9-residue repeats (G-G-X-G-[DN]-D-X-[LVIYF]-X)
F;571,702/Binding site: palmitate (Lys) (covalent) #status predicted

  Query Match             42.1%;  Score 1955.5;  DB 1;  Length 1052;
  Best Local Similarity   42.6%;  Pred. No. 9.1e-93;
  Matches  425;  Conservative  166;  Mismatches  287;  Indels  119;  Gaps   19;

Qy          8 KSNIQAGLNSTKSGLK---------------------NLYLAIPKDYDPQKGGTLNDFI 45

Db         19 KRQAKKGYDVTKNGLQYGVSQAKLQALAAGKAVQKYGNKLVLVIPKEYDGSVGNGFFDLV 78
```

```
Qy         46 KAADELGIARLAEEPNHTETAKKSVDTVNQFLSLTQTGIAISATKLEKFLQKHSTNKLAK 105
              |||:||||        | | | ||: | :||| ||: |: : | :|::||||||
Db         79 KAAEELGIQVKYVNRNELEVAHKSLGTADQFLGLTERGLTLFAPQLDQFLQKHSKISNVV 138

Qy        106 GLDSVENIDRKLGKASNVLSTLSSFLGTALAGIELDSLIKKGDAAPDALAKASIDLINEI 165
              | : : :  || |:  ::| : | ||| |||| |:  |   |  :  ||:| : | :|:
Db        139 GSSTGDAVS-KLAKSQTIISGIQSVLGTVLAGINLNEAIISGGSELE-LAEAGVSLASEL 196

Qy        166 IGNLSQSTQTIEAFSSQLAKLGSTISQAKGFSNIGNKLQNLN---FSKTNLGLEIITGLL 222
              : |::: | ||:||::|:   |  :  |||  :| :|||::    ||| |||:||: ||
Db        197 VSNIAKGTTTIDAFTTQIQNFGKLVENAKGLGGVGRQLQNISGSALSKTGLGLDIISSLL 256

Qy        223 SGISAGFALADKNASTGKKVAAGFELSNQVIGNVTKAISSYVLAQRVAAGLSTTGAVAAL 282
              ||::| ||||:||||  |||||||||||| :|||:|||:||||:|||||||  |||
Db        257 SGVTASFALANKNASTSTKVAAGFELSNQVIGGITKAVSSYILAQRLAAGLSTTGPAAAL 316

Qy        283 ITSSIMLAISPLAFMNAADKFNHANALDEFAKQFRKFGYDGDHLLAEYQRGVGTIEASLT 342
              | ||| |||||||:  || ||  :  :  |||::|:| ||||| ||:|:     |||:||:|
Db        317 IASSISLAISPLAFLRVADNFNRSKEIGEFAERFKKLGYDGDKLLSEFYHEAGTIDASIT 376

Qy        343 TISTALGAVSAGVSAAAVGSAVGTPIALLVAGVTGLISGILEASKQAMFESVANRLQGKI 402
              |||||| |::||  :||: |: || || ||| |:||||||||| ||| | : ||::: ||
Db        377 TISTALSAIAAGTAAASAGALVGAPITLLVTGITGLISGILEFSKQPMLDHVASKIGNKI 436

Qy        403 LEWEKQNGGQNYFDKGYDSRYAAYLANNLKFLSELNKELEAERVIAITQQRWDNNIGELA 462
               ||||:  |:|||: |||:|: |:| ::     ||  ||: | || : |||||| |||||
Db        437 DEWEKKY-GKNYFENGYDARHKAFLEDSFSLLSSFNKQYETERAVLITQQRWDEYIGELA 495

Qy        463 GITKLGERIKSGKAYADAFEDGKKVEAG----SNITLDAKTGIIDISNSNGKKTQALHFT 518
              |||  |:::  |||||  | |::|| :|        | :  |   | ||||||  : :  | |
Db        496 GITKGDKLSSGKAYVDYFQEGKLLEKKPDDFSKVVFDPTKGEIDISNS--QTSTLLKFV 553

Qy        519 SPLLTAGTESRERLTNGKYSYINKLKFGRVKNWQVTDG--EASSKLDFSKVIQR------ 570
              :||||  |||||||  ||| || ||          | | :|  :  :  |:: :||
Db        554 TPLLTPGTESRERTQTGKYEYITKLVVKGKDKW-VVNGVKDKGAVYDYTNLIQHAHISSS 612

Qy        571 VAETEGTDEIGLIVNAKAGNDDIFVGQGKMNIDGGDGHDRVFYSKDGGFGNITVDGTSAT 630
              ||  |   |: |: :  |||  :|:  |    | |:|||| |:| |    | : :||| ||
Db        613 VARGEEYREVRLVSHLGNGNDKVFLAAGSAEIHAGEGHDVVYYDKT-DTGLLVIDGTKAT 671

Qy        631 EAGSYTVNRKVARG-DIYHEVVKRQETKVGKRTETIQYRDYELRKVG-YGYQSTDNLKSV 688
               | | |:| |:::   |  ||:| |:  |||| ||::||||||| :  |    :: | | ||
Db        672 EQGRYSVTRELSGATKILREVIKNQKSAVGKREETLEYRDYELTQSGNSNLKAHDELHSV 731

Qy        689 EEVIGSQFNDVFKGSKFNDIFHSGEGDDLLDGGAGDDRLFGGKGNDRLSGDEGD------ 742
              ||:|||   | ||||||  ||||   :|||||:|  |||  |:| ||||  | || |:
Db        732 EEIIGSNQRDEFKGSKFRDIFHGADGDDLLNGNDGDDILYGDKGNDELRGDNGNDQLYGG 791

Qy        743 --------------------------------------------------DLLDGGSGDDV 753
                                                                ||||||  |:|
Db        792 EGNDKLLGGNGNNYLSGGDGNDELQVLGNGFNVLRGGKGDDKLYGSSGSDLLDGGEGNDY 851

Qy        754 LNGGAGNDVYIFRKGDGNDTLYD---GTGNDKLAFADANISDIMIERTKEGIIVKRNDHS 810
              | || |:| |::|   || |:||     :   |||  :| :   :::|:  : :::: |: |
Db        852 LEGGDGSDFYVYRSTSGNHTIYDQGKSSDLDKLYLSDFSFDRLLVEKVDDNLVLRSNESS 911

Qy        811 ---GSINIPRWYITSNLQNYQSNKTDHKIEQLIGKDGSYITSDQIDKILQDKKDGTVITS 867
                 | :  |  |:        : || :|||||:: |:|  |:| :  :      :  :
Db        912 HNNGVLTIKDWF-------KEGNKYNHKIEQIVDKNGRKLTAENLGTYFKNAPKADNLLN 964

Qy        868 QELKKLADENKSQKLSASDIASSLNKLVGSMALFGTA 904
                  |:  |:|:|  :  | : : |:|:: :   ||  |
Db        965 YATKE--DQNES---NLSSLKTELSKIITNAGNFGVA 996
```

RESULT 6
S51784
toxin III - Actinobacillus pleuropneumoniae (serotype 2)
N;Alternate names: RTX-toxin IIIA (ApxIIIA)
C;Species: Actinobacillus pleuropneumoniae
C;Date: 14-Jul-1995 #sequence_revision 15-Nov-1996 #text_change 18-Jun-1999
C;Accession: S51784
R;Chang, Y.F.; Shi, J.; Ma, D.P.; Shin, S.J.; Lein, D.H.
DNA Cell Biol. 12, 351-362, 1993
A;Title: Molecular analysis of the Actinobacillus pleuropneumoniae RTX toxin-III gene cl
A;Reference number: S51783; MUID:93263992; PMID:8494611
A;Accession: S51784
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1049 <CHA>
A;Cross-references: EMBL:L12145; NID:g349605; PIDN:AAA21924.1; PID:g470685
C;Comment: This orgnism causes porcine pleuropneumonia.
C;Genetics:
A;Gene: apxIIIA
C;Function:
A;Description: lyses lung macrophages
C;Superfamily: hemolysin A; hemolysin A homology
C;Keywords: calcium binding; cytolysis; exotoxin; lipoprotein; tandem repeat; thiolester
F;254-803/Domain: hemolysin A homology <HLYA>
F;735-861/Region: 9-residue repeats (G-G-X-G-[DN]-D-X-[LVIYF]-X)
F;571,702/Binding site: palmitate (Lys) (covalent) #status predicted

```
  Query Match             41.4%;  Score 1924;  DB 1;  Length 1049;
  Best Local Similarity   43.0%;  Pred. No. 3.8e-91;
  Matches  428;  Conservative  152;  Mismatches  296;  Indels  120;  Gaps   21;

Qy          8 KSNIQAGLNSTKSGLK---------------------NLYLAIPKDYDPQKGGTLNDFI 45
              |  :: |  : ||:||:                     | | |||:||   |     | :
Db         19 KRQVKKGYDVTKNGLQYGVSQAKLQALAAGKAVQKYGNKLVLVIPKEYDGSVGNGFFDLV 78

Qy         46 KAADELGIARLAEEPNHTETAKKSVDTVNQFLSLTQTGIAISATKLEKFLQKHSTNKLAK 105
              |||:||||        | | | ||: | :||| ||: |: : | :|::||||||
Db         79 KAAEELGIQVKYVNRNELEVAHKSLGTADQFLGLTERGLTLFAPQLDQFLQKHSKISNVV 138

Qy        106 GLDSVENIDRKLGKASNVLSTLSSFLGTALAGIELDSLIKKGDAAPDALAKASIDLINEI 165
              | : : :  || |:  ::| : | ||| |||| |:  |   |  :  ||:| : | :|:
Db        139 GSSTGDAVS-KLAKSQTIISGIQSVLGTVLAGINLNEAIISGGSELE-LAEAGVSLASEL 196

Qy        166 IGNLSQSTQTIEAFSSQLAKLGSTISQAKGFSNIGNKLQNLN---FSKTNLGLEIITGLL 222
              : |::: | ||:||::|:   |      |||  :| :|||::    ||| |||:||: ||
Db        197 VSNIAKGTTTIDAFTTQIQNFGKLAENAKGLGGVGRQLQNISGSALSKTGLGLDIISSLL 256

Qy        223 SGISAGFALADKNASTGKKVAAGFELSNQVIGNVTKAISSYVLAQRVAAGLSTTGAVAAL 282
              ||::  ||| :|||||  |||||||||||| :|||:|||:||||: |||||  |||
Db        257 SGVTRSFALRNKNASTSTKVAAGFELSNQVIGGITKAVSSYILAQRLRAGLSTTGPAAAL 316

Qy        283 ITSSIMLAISPLAFMNAADKFNHANALDEFAKQFRKFGYDGDHLLAEYQRGVGTIEASLT 342
              | ||| |||||||:  || ||  :  :  |||::|:| ||||| ||:|:     |||:||:|
Db        317 IASSISLAISPLAFLRVADNFNRSKEIGEFAERFKKLGYDGDKLLSEFYHEAGTIDASIT 376

Qy        343 TISTALGAVSAGVSAAAVGSAVGTPIALLVAGVTGLISGILEASKQAMFESVANRLQGKI 402
              |||||| |::||  :||: |: || || ||| |:||||||||| ||| | : ||::: ||
Db        377 TISTALSAIAAGTAAASAGALVGAPITLLVTGITGLISGILEFSKQPMLDHVASKIGNKI 436

Qy        403 LEWEKQNGGQNYFDKGYDSRYAAYLANNLKFLSELNKELEAERVIAITQQRWDNNIGELA 462
               ||||:  |:|||: |||:|: |:| ::     ||  ||: | || : |||||| |||||
Db        437 DEWEKKY-GKNYFENGYDARHKAFLEDSFSLLSSFNKQYETERAVLITQQRWDEYIGELA 495

Qy        463 GITKLGERIKSGKAYADAFEDGKKVEAG----SNITLDAKTGIIDISNSNGKKTQALHFT 518
              |||  |:::  |||||  | |::|| :|        | :  |   | ||||||  : :  | |
Db        496 GITKGDKLSSGKAYVDYFQEGKLLEKKPDDFSKVVFDPTKGEIDISNS--QTSTLLKFV 553

Qy        519 SPLLTAGTESRERLTNGKYSYINKLKFGRVKNWQVTDG--EASSKLDFSKVIQR------ 570
              :||||  |||||||  ||| || ||          | | :|  :  :  |:: :||
Db        554 TPLLTPGTESRERTQTGKYEYITKLVVKGKDKW-VVNGVKDKGAVYDYTNLIQHAHISSS 612

Qy        571 VAETEGTDEIGLIVNAKAGNDDIFVGQGKMNIDGGDGHDRVFYSKDGGFGNITVDGTSAT 630
              ||  |   |: |: :  |||  :|:  |    | |:|||| |:| |    | : :||| ||
Db        613 VARGEEYREVRLVSHLGNGNDKVFLAAGSAEIHAGEGHDVVYYDKT-DTGLLVIDGTKAT 671

Qy        631 EAGSYTVNRKVARG-DIYHEVVKRQETKVGKRTETIQYRDYELRKVG-YGYQSTDNLKSV 688
               | | |:| |:::   |  ||:| |:  |||| ||::||||||| :  |    :: | | ||
Db        672 EQGRYSVTRELSGATKILREVIKNQKYAVGKREETLEYRDYELTQSGNSNLKAHDELHSV 731

Qy        689 EEVIGSQFNDVFKGSKFNDIFHSGEGDDLLDGGAGDDRLFGGKGNDRLSGD--------- 739
              || |||   | ||||||  ||||   :|||||:|  |||  |:| ||||  | ||
```

```
Db      732 EE-IGSNQRDEFKGSKFRDIFHGADGDDLLNGNDGDDILYGDKGNDELRGDNGNDQLYGG 790

Qy      740 EGD---------------------------------------------DLLDGGSGDDV 753

Db      791 EGDDKLLGGNGNNYLSGGDGNDELQVLGNGFNVLRGGKGDDKLYGSSGSDLLDGGEGNDY 850

Qy      754 LNGGAGNDVYIFRKGDGNDTLYD---GTGNDKLAFADANISDIMIERTKEGIIVK--RND 808

Db      851 LEGGDGSDFYVYRSTSGNHTIYDQGKASDSDKLYLSDLSFDNILVKRVNDNLEFRSNNNS 910

Qy      809 HSGSINIPRWYITSNLQNYQSNKTDHKIEQLIGKDGSYITSDQIDKILQDKKDGTVITSQ 868

Db      911 NSGVLTIKDWFKGGNSYN-------HKIEQIVDKNGRKLTAGNLGNNFHD----TQQASS 959

Qy      869 ELKKLADENKSQKLSASDIASSLNKLVGSMALFGTA 904

Db      960 LLKNVTQEQNESNLSS--LKTELGKIITNAGNFGVA 993
```

RESULT 7
S10056
hemolysin A - Escherichia coli plasmid pHly152
C;Species: Escherichia coli
C;Date: 19-Mar-1997 #sequence_revision 25-Apr-1997 #text_change 04-Mar-2000
C;Accession: S10056
R;Hess, J.; Wels, W.; Vogel, M.; Goebel, W.
FEMS Microbiol. Lett. 34, 1-11, 1986
A;Title: Nucleotide sequence of a plasmid-encoded hemolysin determinant and its comparis
A;Reference number: S07209
A;Accession: S10056
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1024 <HES>
A;Cross-references: EMBL:M14107
C;Genetics:
A;Genome: plasmid pHly152
C;Superfamily: hemolysin A; hemolysin A homology
C;Keywords: lipoprotein
F;247-792/Domain: hemolysin A homology <HLYA>
F;564,690/Binding site: palmitate (Lys) (covalent) #status predicted

```
  Query Match             39.8%;  Score 1847.5;  DB 2;  Length 1024;
  Best Local Similarity   41.7%;  Pred. No. 3.1e-87;
  Matches  418;  Conservative  155;  Mismatches  313;  Indels  117;  Gaps   22;

Qy        8 KSNIQAGLNSTKSGLKNLYLAIPKDYDPQKGGTLNDFIKAADELGIARLAEEPNHTETAK 67

Db       32 KDALKKAAEQTRNAGNRLILLIPKDYKGQ-GSSLNDLVRTADELGIEVQYDEKNGTAITK 90

Qy       68 KSVDTVNQFLSLTQTGIAISATKLEKFLQKH--STNKLAKGLDSVENIDRKLGKASNVLS 125

Db       91 QVFGTAEKLIGLTERGVTIFAPQLDKLLQKYQKAGNILGGG---AENIGDNLGKAGGILS 147

Qy      126 TLSSFLGTALAGIELDSLIKK----GDAAPDALAKASIDLINEIIGNLSQSTQTIEAFSS 181

Db      148 TFQNFLGTALSSMKIDELIKKQKSGGNVSSSELAKASIELINQLVDTVASLNNNVNSFSQ 207

Qy      182 QLAKLGSTISQAKGFSNIGNKLQNL-NFSKTNLGLEIITGLLSGISAGFALADKNASTGK 240

Db      208 QLNTLGSVLSNTKHLNGVGNKLQNLPNLDNIGAGLDTVSGILSAISASFILSNADADTRT 267

Qy      241 KVAAGFELSNQVIGNVTKAISSYVLAQRVAAGLSTTGAVAALITSSIMLAISPLAFMNAA 300

Db      268 KAAAGVELTTKVLGNVGKGISQYIIAQRAAQGLSTSAAAAGLIASAVTLAISPLSFLSIA 327

Qy      301 DKFNHANALDEFAKQFRKFGYDGDHLLAEYQRGVGTIEASLTTISTALGAVSAGVSAAAV 360

Db      328 DKFKRANKIEEYSQRFKKLGYDGDSLLAAFHKETGAIDASLTTISTVLASVSSGISAAAT 387

Qy      361 GSAVGTPIALLVAGVTGLISGILEASKQAMFESVANRLQGKILEWEKQNGGQNYFDKGYD 420

Db      388 TSLVGAPVSALVGAVTGIISGILEASKQAMFEHVASKMADVIAEWEKKH-GKNYFENGYD 446

Qy      421 SRYAAYLANNLKFLSELNKELEAERVIAITQQRWDNNIGELAGITKLGERIKSGKAYADA 480

Db      447 ARHAAFLEDNFKILSQYNKEYSVERSVLITQQHWDTLIGELAGVTRNGDKTLSGKSYIDY 506

Qy      481 FEDGKKVEAG----SNITLDAKTGIIDISNSNGKKTQALHFTSPLLTAGTESRERLTNGK 536

Db      507 YEEGKRLEKKPDEFQKQVFDPLKGNIDLSDS--KSSTLLKFVTPLLTPGEEIRERRQSGK 564

Qy      537 YSYINKLKFGRVKNWQVTD-GEASSKLDFSKVIQRVAETEGT-DEIGLIVNAKAGNDDIF 594

Db      565 YEYITELLVKGVDKWTVKGVQDKGAVYDYSNLIQHASVGNNQYREIRIESHLGDGDDKVF 624

Qy      595 VGQGKMNIDGGDGHDRVFYSK-DGGFGNITVDGTSATEAGSYTVNRKVARGD--IYHEVV 651

Db      625 LSAGSANIYAGKGHDVVYYDKTDTGY--LTIDGTKATEAGNYTVTR-VLGGDVKVLQEVV 681

Qy      652 KRQETKVGKRTETIQYRDYELRKV-GYGYQSTDNLKSVEEVIGSQFNDVFKGSKFNDIFH 710

Db      682 KEQEVSVGKRTEKTQYRSYEFTHINGKNLTETDNLYSVEELIGTTRADKFFGSKFTDIFH 741

Qy      711 SGEGDDLLDGGAGDDR------------------------------------------ 726

Db      742 GADGDDLIEGNDGNDRLYGDKGNDTLSGGNGDDQLYGGDGNDKLIGVAGNNYLNGGDGDD 801

Qy      727 ------------LFGGKGNDRLSGDEGDDLLDGGSGDDVLNGGAGNDVYIFRKGDGNDT 773

Db      802 ELQVQGNSLAKNVLFGGKGNDKLYGSEGADLLDGGEGDDLLKGGYGNDIYRYLSGYGHHI 861

Qy      774 LYDGTG-NDKLAFADANISDIMIERTKEGIIVKRND-------HSGSINIPRWYITSNLQ 825

Db      862 IDDDGGKEDKLSLADIDFRDVAFKREGNDLIMYKGEGNVLSIGHKNGITFRNWF-----E 916

Qy      826 NYQSNKTDHKIEQLIGKDGSYITSDQIDKILQDKK-----------DGTVITSQ------ 868

Db      917 KESGDISNHEIEQIFDKSGRIITPDSLKKALEYQQRNNKASYVYGNDALAYGSQGDLNPL 976

Qy      869 --ELKKLADENKSQKLSASDIASSLNKLVGSMALFGTA-NSVS 908

Db      977 INEISKIISAAGSFDVKEERTAASLLQLSGNASDFSYGRNSIT 1019
```

RESULT 8
LEECA
hemolysin A - Escherichia coli
C;Species: Escherichia coli
C;Date: 30-Sep-1988 #sequence_revision 30-Sep-1988 #text_change 18-Jun-1999
C;Accession: A24433; I41280
R;Felmlee, T.; Pellett, S.; Welch, R.A.
J. Bacteriol. 163, 94-105, 1985
A;Title: Nucleotide sequence of an Escherichia coli chromosomal hemolysin.
A;Reference number: A24433; MUID:85234404; PMID:3891743
A;Accession: A24433
A;Molecule type: DNA
A;Residues: 1-1023 <FEL>
A;Cross-references: GB:M10133; GB:M12863; NID:g146377; PIDN:AAA23975.1; PID:g146379
A;Experimental source: strain J96, O4 serotype
R;Stanley, P.; Packman, L.C.; Koronakis, V.; Hughes, C.
Science 266, 1992-1996, 1994
A;Title: Fatty acylation of two internal lysine residues required for the toxic activity
A;Reference number: A55387; MUID:95099325; PMID:7801126
A;Contents: annotation; lysine palmitoylation
A;Note: lysine modification is performed by the hlyC gene product
R;Haertlein, M.; Schiessl, S.; Wagner, W.; Rdest, U.; Kreft, J.; Goebel, W.
J. Cell Biol. 22, 87-97, 1983
A;Title: Transport of hemolysin by Escherichia coli.
A;Reference number: I41280
A;Accession: I41280
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1,'T',3,'V',5,'T',7-44 <RES>
A;Cross-references: GB:M29173; NID:g146337; PIDN:AAA23957.1; PID:g146338
C;Genetics:
A;Gene: hlyA

```
C;Function:
A;Description: attacks blood cell membranes and causes cell lysis
C;Superfamily: hemolysin A; hemolysin A homology
C;Keywords: calcium binding; cytolysis; exotoxin; hemolysis; lipoprotein; tandem repeat;
F;246-791/Domain: hemolysin A homology <HLYA>
F;723-851/Region: 9-residue repeats (G-G-X-G-[DN]-D-X-[LVIYF]-X)
F;563,689/Binding site: palmitate (Lys) (covalent) #status experimental

  Query Match             39.7%;  Score 1845;  DB 1;  Length 1023;
  Best Local Similarity   41.8%;  Pred. No. 4.2e-87;
  Matches  421;  Conservative  156;  Mismatches  311;  Indels  118;  Gaps   23;

Qy          8 KSNIQAGLNSTKSGLKNLYLAIPKDYDPQKGGTLNDFIKAADELGIARLAEEPNHTETAK 67
              |  ::      |::    | | |||||   | | :|||  :: ||||||      :| | |   |
Db         32 KDALKKAAEQTRNAGNRLILLIPKDYKGQ-GSSLNDLVRTADELGIEVQYDEKNGTAITK 90

Qy         68 KSVDTVNQFLSLTQTGIAISATKLEKFLQKH--STNKLAKGLDSVENIDRKLGKASNVLS 125
              :    |  : : ||: |: | |  :|:| |||:  : |||       | |||    ||| :|||
Db         91 QVFGTAEKLIGLTERGVTIFAPQLDKLLQKYQKAGNKLG---GSAENIGDNLGKAGSVLS 147

Qy        126 TLSSFLGTALAGIELDSLIKK----GDAAPDALAKASIDLINEIIGNLSQSTQTIEAFSS 181
              |  :||||||: :::| ||||       |: :    |||||:|||::: : :     : :||
Db        148 TFQNFLGTALSSMKIDELIKKQKSGGNVSSSELAKASIELINQLV-DTAASLNNVNSFSQ 206

Qy        182 QLAKLGSTISQAKGFSNIGNKLQNL-NFSKTNLGLEIITGLLSGISAGFALADKNASTGK 240
              || | |||| :|  |   :|||||||  |  : ::|||   |||  | | |  ::  :|  ||
Db        207 QLNKLGSVLSNTKHLNGVGNKLQNLPNLDNIGAGLDTVSGILSAISASFILSNADADTGT 266

Qy        241 KVAAGFELSNQVIGNVTKAISSYVLAQRVAAGLSTTGAVAALITSSIMLAISPLAFMNAA 300
              | ||| ||: :|:||| | || |::||| | ||||: | | || |  : ||||||:|:: |
Db        267 KAAAGVELTTKVLGNVGKGISQYIIAQRAAQGLSTSAAAAGLIASVVTLAISPLSFLSIA 326

Qy        301 DKFNHANALDEFAKQFRKFGYDGDHLLAEYQRGVGTIEASLTTISTALGAVSAGVSAAAV 360
              |||  ||  ::|::::|:| |||||  |||  : :   |  |:||||  |||  | :||:|:||||
Db        327 DKFKRANKIEEYSQRFKKLGYDGDSLLAAFHKETGAIDASLTRISTVLASVSSGISAAAT 386

Qy        361 GSAVGTPIALLVAGVTGLISGILEASKQAMFESVANRLQGKILEWEKQNGGQNYFDKGYD 420
               | || |::  ||   |||:||||||||||||| ||:::     | ||||::  |:|||: |||
Db        387 TSLVGAPVSALVGAVTGIISGILEASKQAMFEHVASKMADVIAEWEKKH-GKNYFENGYD 445

Qy        421 SRYAAYLANNLKFLSELNKELEAERVIAITQQRWDNNIGELAGITKLGERIKSGKAYADA 480
              :|:||:| :| | |: |||   ||  : ||||  ||    |||||:|:  |::   |||:| |
Db        446 ARHAAFLEDNFKILSQYNKEYSVERSVLITQQHWDTLIGELAGVTRNGDKTLSGKSYIDY 505

Qy        481 FEDGKKVEAG----SNITLDAKTGIIDISNSNGKKTQALHFTSPLLTAGTESRERLTNGK 536
              :|:||::|         |       | ||:|:|   |  :  |  | :||||  | | |||  :||
Db        506 YEEGKRLEKKPDEFQKQVFDPLKGNIDLSDS--KSSTLLKFVTPLLTPGEEIRERRQSGK 563

Qy        537 YSYINKLKFGRVKNWQVTD-GEASSKLDFSKVIQRVAETEGT-DEIGLIVNAKAGNDDIF 594
              | || :|      |   | |        :    |   |:|  :||    :         || :  :     |:| :|
Db        564 YEYITELLVKGVDKWTVKGVQDKGSVYDYSNLIQHASVGNNQYREIRIESHLGDGDDKVF 623

Qy        595 VGQGKMNIDGGDGHDRVFVYSK-DGGFGNITVDGTSATEAGSYTVNRKVARGD--IYHEVV 651
              :   |  ||   | ||  |:|  | | |:   :|:|||  |||||:||| |  |  ||   :   |||
Db        624 LSAGSANIYAGKGHDVVYYDKTDTGY--LTIDGTKATEAGNYTVTR-VLGGDVKVLQEVV 680

Qy        652 KRQETKVGKRTETIQYRDYELRKV-GYGYQSTDNLKSVEEVIGSQFNDVFKGSKFNDIFH 710
              | ||  ||||||  |||  |||            ||||  ||||:||:     |  | ||||  ||||
Db        681 KEQEVSVGKRTEKTQYRSYEFTHINGKNLTETDNLYSVEELIGTTRADKFFGSKFADIFH 740

Qy        711 SGEGDDLLDGGAGDDRLF----------------------------------------- 728
               :|||  ::|    |:|||:
Db        741 GADGDDHIEGNDGNDRLYGDKGNDTLSGGNGDDQLYGGDGNDKLIGGAGNNYLNGGDGDD 800

Qy        729 ---------------GGKGNDRLSGDEGDDLLDGGSGDDVLNGGAGNDVYIFRKGDGNDT 773
                             ||||||:|  |  ||  ||||||  |:|:|  || |||:|  :   |  |:
Db        801 ELQVQGNSLAKNVLSGGKGNDKLYGSEGADLLDGGEGNDLLKGGYGNDIYRYLSGYGHHI 860

Qy        774 LYDGTG-NDKLAFADANISDIMIERTKEGIIVKRND-------HSGSINIPRWYITSNLQ 825
               :  |    |  :|||: || :  |:    |      :|: : :          |      |       |:     :
Db        861 IDDDGGKDDKLSLADIDFRDVAFRREGNDLIMYKAEGNVLSIGHKNGITFKNWF-----E 915

Qy        826 NYQSNKTDHKIEQLIGKDGSYITSDQIDKILQDKK-----------DGTVITSQ------ 868
```

```
                  : ::|:|||:  |||   || |  :  |  |: ::                    |       ||
Db        916 KESGDISNHQIEQIFDKDGRVITPDSLKKALEYQQSNNKASYVYGNDALAYGSQGNLNPL 975

Qy        869 --ELKKLADENKSQKLSASDIASSLNKLVGSMALFGTA-NSVSSNA 911
                |: |:        :    :           |:|| :| |: :  |       ||::   |
Db        976 INEISKIISAAGNFDVKEERAAASLLQLSGNASDFSYGRNSITLTA 1021

RESULT 9
A37205
leukotoxin A - Actinobacillus actinomycetemcomitans
C;Species: Actinobacillus actinomycetemcomitans
C;Date: 16-Sep-1992 #sequence_revision 01-Nov-1996 #text_change 31-Mar-2000
C;Accession: A37205; A60768; B34345; A32276; PH0267; PH0266; S17284
R;Kraig, E.; Dailey, T.; Kolodrubetz, D.
Infect. Immun. 58, 920-929, 1990
A;Title: Nucleotide sequence of the leukotoxin gene from Actinobacillus actinomycetemcom
A;Reference number: A37205; MUID:90202154; PMID:2318535
A;Accession: A37205
A;Molecule type: DNA
A;Residues: 1-1055 <KRA>
A;Cross-references: GB:X16829; NID:g38643; PIDN:CAA34731.1; PID:g38645
A;Note: the authors present evidence that the nucleotide sequence is correct in the vicir
sing nucleotide
R;Kolodrubetz, D.; Dailey, T.; Ebersole, J.; Kraig, E.
Infect. Immun. 57, 1465-1469, 1989
A;Title: Cloning and expression of the leukotoxin gene from Actinobacillus actinomycetemc
A;Reference number: A60768; MUID:89212893; PMID:2707855
A;Accession: A60768
A;Status: nucleic acid sequence not shown; not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 297-309,'Y',311-364;434-440,'KC',443-474,'H',476-489,'S',491-493,'VLK',497-49
A;Note: this sequence is revised in reference A37205
R;Lally, E.T.; Golub, E.E.; Kieba, I.R.; Taichman, N.S.; Rosenbloom, J.; Rosenbloom, J.C
J. Biol. Chem. 264, 15451-15456, 1989
A;Title: Analysis of the Actinobacillus actinomycetemcomitans leukotoxin gene. Delineatic
A;Reference number: A34345; MUID:89359382; PMID:2670940
A;Accession: B34345
A;Molecule type: DNA
A;Residues: 1-239,'Y',241-259,'H',261-335,'A',337-415,'S',417-438,'S',440-723,'N',725-92
PVKRYPLRHCRPIITLTLTQIR' <LAL>
A;Cross-references: GB:M27933
R;Lally, E.T.; Kieba, I.R.; Demuth, D.R.; Rosenbloom, J.; Golub, E.E.; Taichman, N.S.; G
Biochem. Biophys. Res. Commun. 159, 256-262, 1989
A;Title: Identification and expression of the Actinobacillus actinomycetemcomitans leukot
A;Reference number: A32276; MUID:89165863; PMID:2647082
A;Accession: A32276
A;Status: nucleic acid sequence not shown; not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 430-438,'S',440-476,'R',478-506,'RVRS',511,'QSIAINSLNTD',523-541,'I',543,'PQ
A;Note: this sequence is revised in reference A34345
R;Ohta, H.; Miyagi, A.; Kato, K.; Fukui, K.
submitted to JIPID, July 1995
A;Description: Modulation of leukotoxin production by growth rate and bicarbonate in a t
A;Reference number: PH0266
A;Accession: PH0267
A;Molecule type: protein
A;Residues: 17-42 <OHT>
A;Experimental source: strain 301-b
A;Accession: PH0266
A;Molecule type: protein
A;Residues: 2-6,'L',8-26 <OH2>
C;Genetics:
A;Gene: ltxA
C;Function:
A;Description: lyses human polymorphonuclear lymphocytes and monocytes
C;Superfamily: hemolysin A; hemolysin A homology
C;Keywords: cytolysis; hemolysis; lipoprotein; periplasmic space; tandem repeat; thioles
F;245-790/Domain: hemolysin A homology <HLYA>
F;731-847/Region: 9-residue repeats (G-G-X-G-[DN]-D-X-[LVIYF]-X)
F;562,687/Binding site: palmitate (Lys) (covalent) #status predicted

  Query Match             39.1%;  Score 1816;  DB 1;  Length 1055;
```

```
 Best Local Similarity   40.2%;  Pred. No. 1.4e-85;
 Matches  413;  Conservative  185;  Mismatches  299;  Indels  130;  Gaps   23;

Qy          3 NINVIKSNIQAGLNSTKSGLKNLYLAIPKDYDPQKGGTLNDFIKAADELGIARLAEEPNH 62
Db         27 NIDAAKEQLQKALDKLGKTGKKLTLYIPKNY--KKGNGLTALIKAAQKLGIEVYHEGKDG 84

Qy         63 TETAKKSVDTVNQFLSLTQTGIAISATKLEKFLQ--KHSTNKLAKGLDSVENIDRKLGKA 120
Db         85 PALTNGILNTGKKLLGLTERGLTLFAPELDKWIQGNKHLSNSVG----STGNLTKAIDKV 140

Qy        121 SNVLSTLSSFLGTALAGIELDSLIK---KGDAAPDA-LAKASIDLINEIIGNLSQSTQTI 176
Db        141 QSVLGTLQAFLNTAFSGMDLDALIKARQNGKNVTDVQLAKASLNLINELIGTISSITNNV 200

Qy        177 EAFSSQLAKLGSTISQAKGFSNIGNKLQNL----NFSKTNLGLEIITGLLSGISAGFALA 232
Db        201 DTFSKQLNKLGEALGQVKHFGSFGDKLKNLPKLGNLGK---GLGALSGVLSAISAALLLA 257

Qy        233 DKNASTGKKVAAGFELSNQVIGNVTKAISSYVLAQRVAAGLSTTGAVAALITSSIMLAIS 292
Db        258 NKDADTATKAAAAAELTNKVLGNIGKAITQYLIAQRAAAGLSTTGPVAGLIASVVSLAIS 317

Qy        293 PLAFMNAADKFNHANALDEFAKQFRKFGYDGDHLLAEYQRGVGTIEASLTTISTALGAVS 352
Db        318 PLSFLGIAKQFDRARMLEEYSKRFKKFGYNGDSLLGQFYKNTGIADAAITTINTVLSAIA 377

Qy        353 AGVSAAAVGSAVGTPIALLVAGVTGLISGILEASKQAMFESVANRLQGKILEWEKQNGGQ 412
Db        378 AGVGAASAGSLVGAPIGLLVSAITSLISGILDASKQAVFEHIANQLADKIKAWENKY-GK 436

Qy        413 NYFDKGYDSRYAAYLANNLKFLSELNKELEAERVIAITQQRWDNNIGELAGITKLGERIK 472
Db        437 NYFENGYDARHSAFLEDSLKLFNELREKYKTENILSITQQGWDQRIGELAGITRNGDRIQ 496

Qy        473 SGKAYADAFEDGKKVEAGSN----ITLDAKTGIIDISNSNGKKTQALHFTSPLLTAGTES 528
Db        497 SGKAYVDYLKKGEELAKHSDKFTKQILDPIKGNIDLSGIKGSTT--LTFLNPLLTAGKEE 554

Qy        529 RERLTNGKYSYINKLKFGRVKNWQVTDGEASSKL-DFSKVIQR-VAETEGTDEIGLIVNA 586
Db        555 RKTRQSGKYEFITELKVKGRTDWKVKGVPNSNGVYDFSNLIQHAVTRDNKVLEARLIANL 614

Qy        587 KAGNDDIFVGQGKMNIDGGDGHDRVFYSKDGGFGNITVDGTSATEAGSYTVNRKVARGDI 646
Db        615 GAKDDYVFVGSGSTIVNAGDGYDVVDYSK-GRTGALTIDGRNATKAGQYKVERDLSGTQV 673

Qy        647 YHEVVKRQETKVGKRTETIQYRDYEL--RKVGYGYQSTDNLKSVEEVIGSQFNDVFKGSK 704
Db        674 LQETVSKQETKRGKVTDLLEYRNYKLDYYYTNKGFKAHDELNSVEEIIGSTLRDKFYGSK 733

Qy        705 FNDIFHSGEGDDL--------------------------LDGGAGDDRLF--------- 728
Db        734 FNDVFHGHDGDDLIYGYDGDDRLYGDNGNDEIHGGQGNDKLYGGAGNDRLFGEYGNNYLD 793

Qy        729 -----------------GGKGNDRLSGDEGDDLLDGGSGDDVLNGGAGNDVYIFRKGDG 770
Db        794 GGEGDDHLEGGNGSDILRGGSGNDKLFGNQGDDLLDGGEGDDQLAGGEGNDIYVYRKEYG 853

Qy        771 NDTLYDGTGN-DKLAFADANISDIMIERTKEGIIVKRNDHSGSINIPRWYITSN----LQ 825
Db        854 HHTITEHSGDKDKLSLANINLKDVSFERNGNDLLLKTNNRT-AVTFKGWFSKPNSSAGLD 912

Qy        826 NYQ-------------------------SNKTDHKIEQLIGKDGSYITSDQIDKILQDK 859
Db        913 EYQRKLLEYAPEKDRARLKRQFELQRGKVDKSLNNKVEEIIGKDGERITSQDIDNLFDKS 972

Qy        860 KDGTVITSQELKKL-ADENKSQKLSASDIASS-----------LNKLVGSMALFGTA-NS 906
Db        973 GNKKTISPQELAGLIKNKGKSSSLMSSSRSSSMLTQKSGLSNDISRIISATSGFGSSGKA 1032

Qy        907 VSSNALQ 913
Db       1033 LSASPLQ 1039

RESULT 10
I39643
RTX-toxin I - Actinobacillus pleuropneumoniae
N;Alternate names: hemolysin ApxI
C;Species: Actinobacillus pleuropneumoniae
C;Date: 19-Jul-1996 #sequence_revision 08-Nov-1996 #text_change 18-Jun-1999
C;Accession: I39643; S18769; I39645; S60732; S35781
R;Jansen, R.; Briaire, J.; Kamp, E.M.; Gielkens, A.L.; Smits, M.A.
Infect. Immun. 61, 3688-3695, 1993
A;Title: Structural analysis of the Actinobacillus pleuropneumoniae-RTX-toxin I (ApxI) o
A;Reference number: I39641; MUID:93366425; PMID:8359891
A;Accession: I39643
A;Status: nucleic acid sequence not shown; translation not shown; translated from GB/EMB
A;Molecule type: DNA
A;Residues: 1-1022 <RES>
A;Cross-references: EMBL:X73117; NID:g312897; PIDN:CAA51548.1; PID:g312899
R;Frey, J.; Meier, R.; Gygi, D.; Nicolet, J.
Infect. Immun. 59, 3026-3032, 1991
A;Title: Nucleotide sequence of the hemolysin I gene from Actinobacillus pleuropneumonia
A;Reference number: S18769; MUID:91348845; PMID:1879928
A;Accession: S18769
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-209,'AMPYLTLA',218-373,'R',375-561,'Q',563-686,'TC',688-1022 <FRE>
A;Cross-references: EMBL:X52899; NID:g38949; PIDN:CAA37081.1; PID:g38950
R;Frey, J.; Haldimann, A.; Nicolet, J.; Boffini, A.; Prentki, P.
Gene 142, 97-102, 1994
A;Title: Sequence analysis and transcription of the apxI operon (hemolysin I) from Actin
A;Reference number: I39644; MUID:94237497; PMID:8181764
A;Accession: I39645
A;Status: nucleic acid sequence not shown; translation not shown; translated from GB/EMB
A;Molecule type: DNA
A;Residues: 1-209,'AMPYLTLA',218-373,'R',375-561,'Q',563-686,'TC',688-1022 <RE2>
A;Cross-references: EMBL:X68595; NID:g505568; PIDN:CAA48586.1; PID:g505570
R;Tascon, R.I.; Vazquez-Boland, J.A.; Gutierrez-Martin, C.B.; Rodriguez-Barbosa, I.; Rod
Mol. Microbiol. 14, 207-216, 1994
A;Title: The RTX haemolysins ApxI and ApxII are major virulence factors of the swine pat
A;Reference number: S60731; MUID:95131743; PMID:7830567
A;Accession: S60732
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 604-685 <TAS>
C;Comment: This organism causes porcine pleuropneumonia.
C;Genetics:
A;Gene: apxIA
C;Function:
A;Description: attacks blood cell membranes and causes cell lysis
C;Superfamily: hemolysin A; hemolysin A homology
C;Keywords: calcium binding; cytolysis; exotoxin; hemolysis; lipoprotein; tandem repeat;
F;243-789/Domain: hemolysin A homology <HLYA>
F;721-847/Region: 9-residue repeats (G-G-X-G-[DN]-D-X-[LVIYF]-X)
F;560,686/Binding site: palmitate (Lys) (covalent) #status predicted

  Query Match             37.9%;  Score 1763.5;  DB 1;  Length 1022;
  Best Local Similarity   40.0%;  Pred. No. 6.5e-83;
  Matches  400;  Conservative  181;  Mismatches  309;  Indels  109;  Gaps   26;

Qy          8 KSNIQAGLNSTKSGL-------KNLYLAIPKDYDPQKGGTLNDFIKAADELGIARLAEEP 60
Db         21 KSAAKSGAGALKNGLGQVKQAGQKLILYIPKDYQASTGSSLNDLVKAAEALGIEVHRSEK 80

Qy         61 NHTETAKKSVDTVNQFLSLTQTGIAISATKLEKFLQKHSTNKLAKGL-DSVENIDRKLGK 119
Db         81 NGTALAKELFGTTEKLLGFSERGIALFAPQFDKLLNKN--QKLSKSLGGSSEALGQRLNK 138

Qy        120 ASNVLSTLSSFLGTALAGIELDSLIKK----GDAAPDALAKASIDLINEIIGNLSQSTQT 175
Db        139 TQTALSALQSFLGTAIAGMDLDSLLRRRRNGEDVSGSELAKAGVDLAAQLVDNIASATGT 198

Qy        176 IEAFSSQLAKLGSTISQAKGFSNIGNKLQNL-NFSKTNLGLEIITGLLSGISAGFALADK 234
```

```
Db     199 VDAFAEQLGKLGNALSNTR-LSGLASKLNNLPDLSLAGPGFDAVSGILSVVSASFILSNK 257

Qy     235 NASTGKKVAAGFELSNQVIGNVTKAISSYVLAQRVAAGLSTTGAVAALITSSIMLAISPL 294
Db     258 DADAGTKAAAGIEISTKILGNIGKAVSQYIIAQRVAAGLSTTAATGGLIGSVVALAISPL 317

Qy     295 AFMNAADKFNHANALDEFAKQFRKFGYDGDHLLAEYQRGVGTIEASLTTISTALGAVSAG 354
Db     318 SFLNVADKFERAKQLEQYSERFKKFGYEGDSLLASFYRETGAIEAALTTINSVLSAASAG 377

Qy     355 VSAAAVGSAVGTPIALLVAGVTGLISGILEASKQAMFESVANRLQGKILEWEKQNGGQNY 414
Db     378 VGAAATGSLVGAPVAALVSAITGIISGILDASKQAIFERVATKLANKIDEWEKKH-GKNY 436

Qy     415 FDKGYDSRYAAYLANNLKFLSELNKELEAERVIAITQQRWDNNIGELAGITKLGERIKSG 474
Db     437 FENGYDARHSAFLEDTFELLSQYNKEYSVERVVAITQQRWDVNIGELAGITRKGSDTKSG 496

Qy     475 KAYADAFEDGKKVEAG----SNITLDAKTGIIDISNSNGKKTQALHFTSPLLTAGTESRE 530
Db     497 KAYVDFFEEGKLLEKEPDRFDKKVFDPLEGKIDLSSIN--KTTLLKFVTPVFTAGEEIRE 554

Qy     531 RLTNGKYSYINKLKFGRVKNWQVTDGEASSKL-DFSKVIQRVAETEGTDEIGLIVNAKAG 589
Db     555 RKQTGKYEYMTELFVKGKEKWVVTGVQSHNAIYDYTNLIQLAIDKKG-EKRQVTIESHLG 613

Qy     590 --NDDIFVGQGKMNIDGGDGHDRVFYSK-DGGFGNITVDGTSATEAGSYTVNRKV-ARGD 645
Db     614 EKNDRIYLSSGSSIVYAGNGHDVAYYDKTDTGY--LTFDGQSAQKAGEYIVTKELKADVK 671

Qy     646 IYHEVVKRQETKVGKRTETIQYRDYELR--KVGYGYQSTDNLKSVEEVIGSQFNDVFKGS 703
Db     672 VLKEVVKTQDISVGKRSEKLEYRDYELSPFELGNGIRAKDELHSVEEIIGSNRKDKFFGS 731

Qy     704 KFNDIFHSGEGD-----------------DLLDGGAGDDRLFGGKGNDRLSGDEGDDLL 745
Db     732 RFTDIFHGAKGDDEIYGNDGHDILYGDDGNDVIHGGDGNDHLVGGNGNDRLIGGKGNNFL 791

Qy     746 DGGSGDD----------VLNGGAGNDVY-------IFRKGDGNDTLYDGTGND------- 781
Db     792 NGGDGDDELQVFEGQYNVLLGGAGNDILYGSDGTNLFDGGVGNDKIYGGLGKDIYRYSKE 851

Qy     782 -------------------KLAFADANI----SDIMIERTKEGIIVKRNDHSG-SINIPR 817
Db     852 YGRHIIIEKGGDDDTLLLSDLSFKDVGFIRIGDDLLVNKRIGGTLYYHEDYNGNALTIKD 911

Qy     818 WYITSNLQNYQSNKTDHKIEQLIGKDGSYITSDQIDKILQDKKDGTVITSQELKKLADEN 877
Db     912 WF--KEGKEGQNN----KIEKIVDKDGAYVLSQYLTELTAPGRGINYFNGLEEKLYYGEG 965

Qy     878 KSQKLSASDIASSLNKLVGSMALF-GTANSVSSNALQPI 915
Db     966 YN---ALPQLRKDIEQIISSTGAFTGDHGKVSVGSGGPL 1001
```

RESULT 11
T00227
hemolysin A toxin protein - Escherichia coli plasmid pO157
C;Species: Escherichia coli
C;Date: 01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change 21-Jul-2000
C;Accession: T00227; T42148
R;Makino, K.; Ishii, K.; Yasunaga, T.; Hattori, M.; Yokoyama, K.; Yatsudo, H.C.; Kubota, S.; Shinagawa, H.
DNA Res. 5, 1-9, 1998
A;Title: Complete nucleotide sequences of 93-kb and 3.3-kb plasmids of an enterohemorrha
A;Reference number: Z14127; MUID:98290540; PMID:9628576
A;Accession: T00227
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-998 <MAK>
A;Cross-references: EMBL:AB011549; NID:g4589740; PIDN:BAA31774.1; PID:g3337015
A;Experimental source: strain EHEC O157:H7, substrain RIMD 0509952
R;Burland, V.; Shao, Y.; Perna, N.T.; Plunkett, G.; Sofia, H.J.; Blattner, F.R.
Nucleic Acids Res. 26, 4196-4204, 1998
A;Title: The complete DNA sequence and analysis of the large virulence plasmid of Escher
A;Reference number: Z22068; MUID:98391744; PMID:9722640
A;Accession: T42148
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-998 <BUR>
A;Cross-references: EMBL:AF074613; PIDN:AAC70116.1
A;Experimental source: strain EDL933; serotype O157:H7
C;Genetics:
A;Gene: hlyA
A;Genome: plasmid pO157
C;Superfamily: hemolysin A; hemolysin A homology
C;Keywords: cytolysis; hemolysis; lipoprotein; toxin
F;233-776/Domain: hemolysin A homology <HLYA>
F;550,675/Binding site: palmitate (Lys) (covalent) #status predicted

```
  Query Match             37.6%;  Score 1748;  DB 2;  Length 998;
  Best Local Similarity   40.0%;  Pred. No. 3.9e-82;
  Matches  391;  Conservative  177;  Mismatches  308;  Indels  102;  Gaps   21;

Qy       8 KSNIQAGLNSTKSGLKNLYLAIPKDYDPQKGGTLNDFIKAADELGIARLAEEPNHTETAK 67
Db      18 KSAFNTASSSVRSAGKKLILLIPDNYEAQ-GVGINELVKAADELGIEIHRTERDDTAIAN 76

Qy      68 KSVDTVNQFLSLTQTGIAISATKLEKFLQKHSTNKLAKGL-DSVENIDRKLGKASNVLST 126
Db      77 QFFGAAEKVVGLTERGVAIFAPQLDKLLQKY--QKVGSKIGGTAENVGNNLGKAGTVLSA 134

Qy     127 LSSFLGTALAGIELDSLIKK----GDAAPDALAKASIDLINEIIGNLSQSTQTIEAFSSQ 182
Db     135 LQNFTGIALSGMALDELLRKQREGEDISQNDIAKSSIELINQLVDTVSSINSTVDSFSEQ 194

Qy     183 LAKLGSTISQAKGFSNIGNKLQNL-NFSKTNLGLEIITGLLSGISAGFALADKNASTGKK 241
Db     195 LNQLGSFLSSKPRLSSVGGKLQNLPDLGPLGDGLDVVSGILSAVSASFILGNSDAHTGTK 254

Qy     242 VAAGFELSNQVIGNVTKAISSYVLAQRVAAGLSTTGAVAALITSSIMLAISPLAFMNAAD 301
Db     255 AAAGIELTTQVLGNVGKAVSQYILAQRMAQGLSTTAASAGLITSAVMLAISPLSFLAAAD 314

Qy     302 KFNHANALDEFAKQFRKFGYDGDHLLAEYQRGVGTIEASLTTISTALGAVSAGVSAAAVG 361
Db     315 KFERAKQLESYSERFKKLNYEGDALLAAFHKETGAIDAALTTINTVLSSVSAGVSAASSA 374

Qy     362 SAVGTPIALLVAGVTGLISGILEASKQAMFESVANRLQGKILEWEKQNGGQNYFDKGYDS 421
Db     375 SLIGAPISMLVSALTGTISGILEASKQAMFEHVAEKFAARINEWEKEH-GKNYFENGYDA 433

Qy     422 RYAAYLANNLKFLSELNKELEAERVIAITQQRWDNNIGELAGITKLGERIKSGKAYADAF 481
Db     434 RHAAFLEDSLSLLADFSRQHAVERAVAITQQHWDEKIGELAGITRNADRSQSGKAYINYL 493

Qy     482 EDGKKVEAG----SNITLDAKTGIIDISNSNGKKTQALHFTSPLLTAGTESRERLTNGKY 537
Db     494 ENGGLLEAQPKEFTQQVFDPQKGTIDL--STGNVSSVLTFITPTFTPGEEVRERKQSGKY 551

Qy     538 SYINKLKFGRVKNWQVTDGEASSK--LDFSKVIQRVAETEGTDEIGLIVNAKAGNDDIFV 595
Db     552 EYMTSLIVNGKDTWSV-KGIKNHKGVYDYSKLIQFVEKNNKHYQARIISELGDKDDVVYS 610

Qy     596 GQGKMNIDGGDGHDRVFYSKDGGFGNITVDGTSATEAGSYTVNRKVARGD--IYHEVVKR 653
Db     611 GAGSSEVFAGEGYDTVSYNKT-DVGKLTIDATGASKPGEYIVSKNM-YGDVKVLQEVVKE 668

Qy     654 QETKVGKRTETIQYRDYELRKVGYGYQSTDNLKSVEEVIGSQFNDVFKGSKFNDIFHSG- 712
Db     669 QEVSVGKRTEKIQYRDFEFRTGGIPYDVIDNLHSVEELIGGKHDDEFKGGKFNDIFHGAD 728

Qy     713 ----------------EGDDLLDGGAGDDRLFGGKGNDRLSGDEGD------------- 742
Db     729 GNDYIEGNYGNDRLYGDDGDDYISGGQGDDQLFGGSGNDKLSGGDGNNYLTGGSGNDELQ 788
```

```
Qy        743 ----------------------DLLDGGSGDDVLNGGAGNDVYIFRKGDGNDTLYD-GT 778
Db        789 AHGAYNILSGGTGDDKLYGGGGIDLLDGGEGNDYLNGGFGNDIYVYGQNYGHHTIADEGG 848

Qy        779 GNDKLAFADANISDIMIERTKEGIIVKR-----------NDHSGSINIPRWYITSNLQNY 827
Db        849 KGDRLHLSDISFDDIAFKRVGNDLIMNKAINGVLSFNESNDVNG-ITFKNWFAKD----- 902

Qy        828 QSNKTDHKIEQLIGKDGSYITSDQIDKILQDKKDGTVITSQELKKLADENKSQKLSASDI 887
Db        903 ASGADNHLVEVITDKDGREI---KVDKIPHNNNE----RSGYIKASNIASEKNMVNITSV 955

Qy        888 ASSLNKLVGSMALFGTAN 905
Db        956 ANDINKIISSVSGFDSGD 973
```

RESULT 12
I41078
hemolysin - Escherichia coli
C;Species: Escherichia coli
C;Date: 31-May-1996 #sequence_revision 31-May-1996 #text_change 21-Jul-2000
C;Accession: I41078
R;Schmidt, H.; Beutin, L.; Karch, H.
Infect. Immun. 63, 1055-1061, 1995
A;Title: Molecular analysis of the plasmid-encoded hemolysin of Escherichia coli O157:H7
A;Reference number: I41077; MUID:95172699; PMID:7868227
A;Accession: I41078
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-998 <RES>
A;Cross-references: EMBL:X79839; NID:g860924; PIDN:CAA56234.1; PID:g4388764
C;Superfamily: hemolysin A; hemolysin A homology
C;Keywords: lipoprotein
F;233-776/Domain: hemolysin A homology <HLYA>
F;550,675/Binding site: palmitate (Lys) (covalent) #status predicted

```
  Query Match             37.2%;  Score 1729;  DB 2;  Length 998;
  Best Local Similarity   39.8%;  Pred. No. 3.7e-81;
  Matches  389;  Conservative  174;  Mismatches  313;  Indels  102;  Gaps   21;

Qy          8 KSNIQAGLNSTKSGLKNLYLAIPKDYDPQKGGTLNDFIKAADELGIARLAEEPNHTETAK 67
Db         18 KSAFNTASSSVRSAGKKLILLIPDNYEAQ-GVGINELVKAADELGIEIHRTERDDTAIAN 76

Qy         68 KSVDTVNQFLSLTQTGIAISATKLEKFLQKHSTNKLAKGL-DSVENIDRKLGKASNVLST 126
Db         77 QFFGAAEKVVGLTERGVAIFAPQLDKLLQKY--QKVGSKIGGTAENVGNNLGKAGTVLSA 134

Qy        127 LSSFLGTALAGIELDSLIKK----GDAAPDALAKASIDLINEIIGNLSQSTQTIEAFSSQ 182
Db        135 LQNFTGIALSGMALDELLRKQREGEDISQNDIAKSSIELINQLVDTVSSINSTVDSFSEQ 194

Qy        183 LAKLGSTISQAKGFSNIGNKLQNL-NFSKTNLGLEIITGLLSGISAGFALADKNASTGKK 241
Db        195 LNQLGSFLSSKPRLSSVGGKLQNLPDLGPLGDGLDVVSGILSAVSASFILGNSDAHTGTK 254

Qy        242 VAAGFELSNQVIGNVTKAISSYVLAQRVAAGLSTTGAVAALITSSIMLAISPLAFMNAAD 301
Db        255 AAAGIELTTQVLGNVGKAVSQYILAQRMAQGLSTTAASAGLITSAVMLAISPLSFLAAAD 314

Qy        302 KFNHANALDEFAKQFRKFGYDGDHLLAEYQRGVGTIEASLTTISTALGAVSAGVSAAAVG 361
Db        315 KFERAKQLESYSERFKKLNYEGDALLAGFHKETGAIDAGLTTINTVLSSVSAGVSAASSA 374

Qy        362 SAVGTPIALLVAGVTGLISGILEASKQAMFESVANRLQGKILEWEKQNGGQNYFDKGYDS 421
Db        375 SLIGAPISMLVSALTGTISGILEASKQAMFEHVAEKFAARINEWEKEH-GKNYFENGYDA 433

Qy        422 RYAAYLANNLKFLSELNKELEAERVIAITQQRWDNNIGELAGITKLGERIKSGKAYADAF 481
Db        434 RHAAFLEDSLSLLADFSRQHAVERAVAITHQHWDEKIGELAGITRNADRSQSGKPYINYL 493

Qy        482 EDGKKVEAG----SNITLDAKTGIIDISNSNGKKTQALHFTSPLLTAGTESRERLTNGKY 537
Db        494 ENGGLLEAQPKEFTQQVFDPQKGTIDL--STGNVSSVLTFITPTFTPGEEVRERKQSGKY 551

Qy        538 SYINKLKFGRVKNWQVTDGEASSK--LDFSKVIQRVAETEGTDEIGLIVNAKAGNDDIFV 595
Db        552 EYMTSLIVNGKDTWSV-KGIKNHKGVYDYSKLIQFVEKNTKHYQARIISELGDKDDVVYS 610

Qy        596 GQGKMNIDGGDGHDRVFYSKDGGFGNITVDGTSATEAGSYTVNRKVARGD--IYHEVVKR 653
Db        611 GAGSSEVFAGEGYDPVSYNKT-DVGKLTIDATGAPKPGEYIVPKNM-YGDVEVLQEVVKE 668

Qy        654 QETKVGKRTETIQYRDYELRKVGYGYQSTDNLKSVEEVIGSQFNDVFKGSKFNDIFHSG- 712
Db        669 QEVSVGKRTEKIQYRDFEFRTGGIPYDVIDNLHSVEELIGGKHDDEFKGGKFNDIFHGAD 728

Qy        713 -----------------EGDDLLDGGAGDDRLFGGKGNDRLSGDEGD------------- 742
Db        729 GNDYIEGNYGNDRLYGDDGDDYISGGQGDDQLFGGSGNDKLSGGDGNNYLTGGSGNDELQ 788

Qy        743 ----------------------DLLDGGSGDDVLNGGAGNDVYIFRKGDGNDTLYD-GT 778
Db        789 AHGAYNILSGGTGDDKLYGGGGIDLLDGGEGNDYLNGGFGNDIYVYGQNYGHHTIADEGG 848

Qy        779 GNDKLAFADANISDIMIERTKEGIIVKR-----------NDHSGSINIPRWYITSNLQNY 827
Db        849 KGDRLHLSDISFDDIAFKRVGNDLIMNKAINGVLSFNESNDVNG-ITFKNWFAKD----- 902

Qy        828 QSNKTDHKIEQLIGKDGSYITSDQIDKILQDKKDGTVITSQELKKLADENKSQKLSASDI 887
Db        903 ASGADNHLVEVITDKDGREI---KVDKIPHNNNE----RSGYIKASNIASEKNMVNITSV 955

Qy        888 ASSLNKLVGSMALFGTAN 905
Db        956 ANDINKIISSVSGFDSGD 973
```

RESULT 13
S51672
adenylate cyclase hemolysin - Bordetella bronchiseptica
C;Species: Bordetella bronchiseptica
C;Date: 07-May-1995 #sequence_revision 01-Sep-1995 #text_change 16-Jul-1999
C;Accession: S51672
R;Betsou, F.; Sismeiro, O.; Danchin, A.; Guiso, N.
submitted to the EMBL Data Library, September 1994
A;Description: The adenylate cyclase-hemolysin gene from Bordetella bronchiseptica.
A;Reference number: S51672
A;Accession: S51672
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1705 <BET>
A;Cross-references: EMBL:Z37112
C;Superfamily: cyclolysin; calmodulin-sensitive adenylate cyclase catalytic domain homol
C;Keywords: tandem repeat
F;15-328/Domain: calmodulin-sensitive adenylate cyclase catalytic domain homology <ADE>
F;543-1084/Domain: hemolysin A homology <HLYA>

```
  Query Match             17.3%;  Score 805.5;  DB 2;  Length 1705;
  Best Local Similarity   27.7%;  Pred. No. 2.4e-33;
  Matches  254;  Conservative  137;  Mismatches  334;  Indels  193;  Gaps   28;

Qy        106 GLDSVENIDRKLGKASNVLSTLSSFLGTALAGIELDSLI-----KKGDAAP--------- 151
Db        403 GYDSLDGV----GSRSFSLGEVSDMAAVEAAELEMTRQVLHAGARQDDAEPGVSGASAHW 458

Qy        152 ------DALAKASIDLINEIIGNLSQ-----STQT-IEAFSSQLAKLG---------STI 190
Db        459 GQRALQGAQAVAAAQRLVHAIALMTQFGRAGSTNTPQEAASLSAAVFGLGEASSAVAETV 518

Qy        191 S--------QAKGFSNIGNKLQNLNFSKTNLGLEIITGLLSGISAGFALADKNASTGKKV 242
Db        519 SGFFRGSSRWAGGFGVAG-------------GAMALGGGIGAVGAGMSLTD-DAPAGQKA 564
```

```
Qy        243 AAGFELSNQVIGNVTKAISSYVL----AQRVAAGLSTTGAVAALITSSIMLAISPLAFMN 298
Db        565 AAGAEIALQLTGGTVELASSIALALAAARGVTSGLQVAGASAGAAAGALAAALSPMEIYG 624

Qy        299 AADKFNHANALDEFAKQFRKFGYDGDHLLAEYQRGVGTIEASLTTISTALGAVSAGVSAA 358
Db        625 LVQQSHYADQLDKLAQESSAYGYEGDALLAQLYRDKTAAEGAVAGVSAVLSTVGAAVSIA 684

Qy        359 AVGSAVGTPIALLVAGVTGLISGILEASKQAMFESVANRLQGKILEWEKQNGGQNYFDKG 418
Db        685 AAASVVGAPVAVVTSLLTGALNGILRGVQQPIIEKLANDYARKI---DELGGPQAYFEKN 741

Qy        419 YDSRYAAYLANN---LKFLSELNKELEAERVIAITQQRWDNNIGELAGITKLGERIKSGK 475
Db        742 LQARH-EQLANSDGLRKMLADLQAGWNASSVIGVQTTEISKSALELAAITGNADNLKSAD 800

Qy        476 AYADAFEDGKKVEAGSNITLDAKTGIIDISNSNGKKTQALHFTSPLLTAGTESRERLTNG 535
Db        801 VFVDRFIQGERV-AGQPVVLDVAAGGIDIASRKGER-PALTFITPLAAPGEEQRRRTKTG 858

Qy        536 KYSYINKLKF-GRVKNWQVTDGEASSKLDFSKVIQRVAETEGT--DEIGLIVNAKAGNDD 592
Db        859 KSEFTTFVEIVGKQDRWRIRDGAADTTIDLAKVVSQLVDANGVLKHSIKLEVIGGDGDDV 918

Qy        593 IFVGQGKMNIDGGDGHDRVFYSKDGGFGNITVDGTSATEAGSYTVNRKVARGDIYHEVVK 652
Db        919 VLANASRIHYDGGAGTNTVSYAALGRQDSITV----SADGERFNVRKQLNNANVYREGVA 974

Qy        653 RQETKVGKRTETIQYRDYELRKVGYGYQSTDNLKSVEEVIGSQFNDVFKGSKFNDIFHSG 712
Db        975 TQKTAYGKRTENVQYRHVELARVGQ-LVEVDTLEHVQHIIGGAGNDSITGNAHDNFLAGG 1033

Qy        713 EGDDLLDGGAGDDRLFGGKGNDRLSGDEGDDL-----------LDGGSG----------- 750
Db       1034 AGDDRLDGGAGNDTLVGGEGHNTVVGGAGDDVFLQDLGVWSNQLDGGAGVDTVKYNVHQP 1093

Qy        751 ----------------------------------------------------------DDVL 754
Db       1094 SEERLERMGDTGIHADLQKGTVEKWPALNLFSVDHVKNIENLHGSSLNDSIAGDDRDNEL 1153

Qy        755 NGGAGNDVYIFRKGD-------GNDTLYDGTGNDKLAFADANISDIM-----IERTKEGI 802
Db       1154 WGHDGNDTIHGRRGDDILRGGLGLDTLYGEDGNDIFLQDDETVSDDIDGGAGLDTVDYAA 1213

Qy        803 IVKRNDHSGSINIPRWY---ITSNLQNYQSNKTDHK----------IEQLIG---KDGSY 846
Db       1214 MI----HAGNIVAPHEYGFGIEADLSEGWVRKAARRGMDYYDSVRSVENVIGTSMKD-VL 1268

Qy        847 ITSDQIDKILQDKKDGTV 864
Db       1269 IGDAQANTLMGQGGDDTV 1286
```

RESULT 14
OYBRC
cyclolysin - Bordetella pertussis
N;Alternate names: adenylate cyclase precursor; calmodulin-sensitive adenylate cyclase-h
N;Contains: adenylate cyclase (EC 4.6.1.1), calmodulin-sensitive; hemolysin
C;Species: Bordetella pertussis
C;Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 16-Jul-1999
C;Accession: S00893; S14100; S02389
R;Glaser, P.; Ladant, D.; Sezer, O.; Pichot, F.; Ullmann, A.; Danchin, A.
Mol. Microbiol. 2, 19-30, 1988
A;Title: The calmodulin-sensitive adenylate cyclase of Bordetella pertussis: cloning and
A;Reference number: S00893; MUID:88216178; PMID:2897067
A;Accession: S00893
A;Molecule type: DNA
A;Residues: 1-1706 <GLA>
A;Cross-references: EMBL:Y00545; NID:g396665; PIDN:CAA68613.1; PID:g396666
R;Glaser, P.; Sakamoto, H.; Bellalou, J.; Ullmann, A.; Danchin, A.
EMBO J. 7, 3997-4004, 1988
A;Title: Secretion of cyclolysin, the calmodulin-sensitive adenylate cyclase--haemolysin
A;Reference number: S02386; MUID:89091151; PMID:2905265
A;Contents: annotation; identification of adenylate cyclase--hemolysin bifunctional prote
R;Munier, H.; Gilles, A.M.; Glaser, P.; Krin, E.; Danchin, A.; Sarfati, R.; Barzu, O.
Eur. J. Biochem. 196, 469-474, 1991
A;Title: Isolation and characterization of catalytic and calmodulin-binding domains of Bo
A;Reference number: S14100; MUID:91177021; PMID:2007407
A;Accession: S14100
A;Molecule type: protein
A;Residues: 1-78,'M',80,'M',82-97,'M',99-139,'M',141-178,'M',180-399 <MUN>
R;Hackett, M.; Guo, L.; Shabanowitz, J.; Hunt, D.F.; Hewlett, E.L.
Science 266, 433-435, 1994
A;Title: Internal lysine palmitoylation in adenylate cyclase toxin from Bordetella pertus
A;Reference number: A55167; MUID:95025937; PMID:7939682
A;Contents: annotation; lysine palmitoylation
C;Comment: B. pertussis, the etiological agent of whooping cough, disrupts mammalian cell
ation by host cell calmodulin of the adenylate cyclase activity of bacterial cyclolysin,
C;Comment: Adenylate cyclase activity is activated upon binding of calmodulin in the vic
C;Genetics:
A;Gene: cyaA; cya
C;Superfamily: cyclolysin; calmodulin-sensitive adenylate cyclase catalytic domain homolo
C;Keywords: calcium binding; calmodulin binding; cAMP biosynthesis; carbon-oxygen lyase;
F;15-328/Domain: calmodulin-sensitive adenylate cyclase catalytic domain homology <ADE>
F;25-35/Region: calmodulin binding #status predicted
F;59-66/Region: nucleotide binding #status predicted
F;544-1085/Domain: hemolysin A homology <HLYA>
F;1033-1041,1042-1050,1174-1182,1289-1297,1298-1308,1316-1324,1430-1438,1556-1564/Region
F;983/Binding site: palmitate (Lys) (covalent) #status experimental

```
  Query Match             17.3%;  Score 802;  DB 1;  Length 1706;
  Best Local Similarity   27.1%;  Pred. No. 3.7e-33;
  Matches  246;  Conservative  142;  Mismatches  347;  Indels  174;  Gaps   26;

Qy        106 GLDSVENIDRKLGKASNVLSTLSSFLGTALAGIELDSLI-----KKGDAAP--------- 151
Db        403 GYDSLDGV----GSRSFSLGEVSDMAAVEAAELEMTRQVLHAGARQDDAEPGVSGASAHW 458

Qy        152 --------DALAKA-----SIDLINEI----IGNLSQSTQTIEAFSSQLAKLGSTISQ-A 193
Db        459 GQRALQGAQAVAAAQRLVHAIALMTQFGRAGSTNTPQEAASLSAAVFGLGEASSAVAETV 518

Qy        194 KGFSNIGNKLQNLNFSKTNLGLEIITGLLSGISAGFALADKNASTGKKVAAGFELSNQVI 253
Db        519 SGFFR-GSSRWAGGFGVAGGAMALGGGIAAAVGAGMSLTD-DAPAGQKAAAGAEIALQLT 576

Qy        254 GNVTKAISSYVL----AQRVAAGLSTTGAVAALITSSIMLAISPLAFMNAADKFNHANAL 309
Db        577 GGTVELASSIALALAAARGVTSGLQVAGASAGAAAGALAAALSPMEIYGLVQQSHYADQL 636

Qy        310 DEFAKQFRKFGYDGDHLLAEYQRGVGTIEASLTTISTALGAVSAGVSAAAVGSAVGTPIA 369
Db        637 DKLAQESSAYGYEGDALLAQLYRDKTAAEGAVAGVSAVLSTVGAAVSIAAAASVVGAPVA 696

Qy        370 LLVAGVTGLISGILEASKQAMFESVANRLQGKILEWEKQNGGQNYFDKGYDSRYAAYLAN 429
Db        697 VVTSLLTGALNGILRGVQQPIIEKLANDYARKI---DELGGPQAYFEKNLQARH-EQLAN 752

Qy        430 N---LKFLSELNKELEAERVIAITQQRWDNNIGELAGITKLGERIKSGKAYADAFEDGKK 486
Db        753 SDGLRKMLADLQAGWNASSVIGVQTTEISKSALELAAITGNADNLKSVDVFVDRFVQGER 812

Qy        487 VEAGSNITLDAKTGIIDISNSNGKKTQALHFTSPLLTAGTESRERLTNGKYSYINKLKF- 545
Db        813 V-AGQPVVLDVAAGGIDIASRKGER-PALTFITPLAAPGEEQRRRTKTGKSEFTTFVEIV 870

Qy        546 GRVKNWQVTDGEASSKLDFSKVIQRVAETEGT--DEIGLIVNAKAGNDDIFVGQGKMNID 603
Db        871 GKQDRWRIRDGAADTTIDLAKVVSQLVDANGVLKHSIKLDVIGGDGDDVVLANASRIHYD 930

Qy        604 GGDGHDRVFYSKDGGFGNITVDGTSATEAGSYTVNRKVARGDIYHEVVKRQETKVGKRTE 663
Db        931 GGAGTNTVSYAALGRQDSITV----SADGERFNVRKQLNNANVYREGVATQTTAYGKRTE 986

Qy        664 TIQYRDYELRKVGYGYQSTDNLKSVEEVIGSQFNDVFKGSKFNDIFHSGEGDDLLDGGAG 723
```

```
                  :|||  || :||    :  | |: |: :||   ||    |:  ::     | ||| ||||||
Db        987 NVQYRHVELARVGQVVE-VDTLEHVQHIIGGAGNDSITGNAHDNFLAGGSGDDRLDGGAG 1045

Qy        724 DDRLFGGKG-------------------------------------------------- 732
              :| | ||:|
Db       1046 NDTLVGGEGQNTVIGGAGDDVFLQDLGVWSNQLDGGAGVDTVKYNVHQPSEERLERMGDT 1105

Qy        733 ------------------------------------NDRLSGDEGDDLLDGGSGDDVLNG 756
                                                  |||::||: |: |  |  |:| : |
Db       1106 GIHADLQKGTVEKWPALNLFSVDHVKNIENLHGSRLNDRIAGDDQDNELWGHDGNDTIRG 1165

Qy        757 GAGNDVYIFRKGDGNDTLYDGTGNDKLAFADANISDIM-----IERTKEGIIVKRNDHSG 811
                |:|  | | | | ||||  |||      |   :|| :      ::     ::    | |
Db       1166 RGGDD--ILRGGLGLDTLYGEDGNDIFLQDDETVSDDIDGGAGLDTVDYSAMI----HPG 1219

Qy        812 SINIPRWY---ITSNLQNYQSNKTD---------HKIEQLIG---KDGSYITSDQIDKI 855
                |   |    |  ::|         |             :| :||    ||    |     | : :
Db       1220 RIVAPHEYGFGIEADLSREWVRKASALGVDYYDNVRNVENVIGTSMKD-VLIGDAQANTL 1278

Qy        856 LQDKKDGTV 864
              :    | ||
Db       1279 MGQGGDDTV 1287
```

RESULT 15
E81086
iron-regulated protein FrpC NMB1415 [imported] - Neisseria meningitidis (strain MC58 ser
C;Species: Neisseria meningitidis
C;Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-Jan-2001
C;Accession: E81086
R;Tettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A
Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.;
ri, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Masignani, V.; Pizza, M.
Science 287, 1809-1815, 2000
A;Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; Ve
A;Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.
A;Reference number: A81000; MUID:20175755; PMID:10710307
A;Accession: E81086
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1829 <TET>
A;Cross-references: GB:AE002490; GB:AE002098; NID:g7226651; PIDN:AAF41776.1; PID:g722665
A;Experimental source: serogroup B, strain MC58
C;Genetics:
A;Gene: NMB1415

```
  Query Match             9.4%;  Score 437;  DB 2;  Length 1829;
  Best Local Similarity   20.4%;  Pred. No. 2.4e-14;
  Matches  227;  Conservative  175;  Mismatches  365;  Indels  346;  Gaps   46;

Qy          6 VIKSNIQAGLNSTKSGLKNLYLAIPKDYDPQKGGTLNDFIKAADELGIARLAEEPNHTET 65
              :: :||:  |   : |: : :|                  ::|    |  | |
Db        243 IVNNNIKQGNEAFKNEINSL---------------VHDMKAAGKEFG------------ 274

Qy         66 AKKSVDTVN-QFLSLTQTGIAISATKLEKFLQKHSTNKLAKGLDSVENIDRKLGKASNVL 124
                    | :| |: :|||      :|   :    :  :::  :  ||: ::: :    |:   |:: |
Db        275 -----DDLNTQWNNLTQ-----AAEIIYNDIVDNTSQGIEKGVKAIKELSEKMKNAASDL 324

Qy        125 STLSSFLGTALAGIELDSLIKKGDAAPDALAKASIDLINEIIGNLSQSTQTIEAFSSQLA 184
              :                |    :|       :  ||:|: :              :: |    |    |
Db        325 A---------------DGSAEKAKQVVEDLAQAAKEAYENAKSTAEKAAQAAREFFKGLP 369

Qy        185 K------------------------LGSTISQAK------------------GFSNIG 200
                                      |   : : |                     |  :
Db        370 SFKDLAEKFRDLFPNPEGWIDDGHQCLAPWVKETKKRNGKYHVYDPLALDLDGDGIETVA 429

Qy        201 NK-LQNLNFSKTNLGLEIITGLLSGISAGFALADKNASTGKKVAAGFELSNQVIGNVTKA 259
               |       | ||:     || :|        |   :  |  |                :|  |  :
Db        430 TKGFAGSLFDHTNNGIRTATGWVSA-DDGLLVRDLN-------------GNGIIDNGAEL 475

Qy        260 ISSYVLAQRVAAGLSTTGAVAALITSSIMLAISPLAFMNAADKFNHANALDEFAKQFRKF 319
                      ::| |       |||                |   :  |  |:|   :  ::
Db        476 FGDNT---KLADGSFAKHGYAALAELD----------SNGDNIINAADAAFQTLRVWQDL 522

Qy        320 GYDG----DHLLAEYQRGVGTIEASLTTISTALGAVSAGVSAAAVGSAVGTPIALLVAGV 375
                ||    :  |       :  |: ::: :   ::  ||      |  :  |  ||    |       |
Db        523 NQDGISQANELRTLEELGIQSLDLAYKDVNKNLG---NGNTLAQQGSYTKTD------GT 573

Qy        376 TGLISGILEASKQAMFESVANRLQGKILEWEKQNGGQNYFDKG--YDSRYAAYLANNLKF 433
                |   :  :| |:      ::: :| |: |:    :|      |       |     |  | || |: :|
Db        574 TAKMGDLLLAA-----DNLHSRFKDKVELTAEQAKAANLAGIGRLRDLREAAALSGDLAN 628

Qy        434 LSELNKELEA-ERVIAITQ---QRW---DNNIGELA-------------GITKLGERI- 471
              : :        |   |  :|:      :|      |:| |: :                 ||       ::
Db        629 MLKAYSAAETKEAQLALLDNLIHKWAETDSNWGKKSPMRLSTDWTQTANEGIALTPSQVA 688

Qy        472 -----------KSGKAYADAFEDGKKVEAGSNITLDAKTG---------------IIDI 504
                          ||  ||  |   |            ||| ||                   |: :
Db        689 QLKKNALVSLSDKAKAAIDAARDRIAV-------LDAYTGQDSNTLYYMSEEDALNIVKV 741

Qy        505 SNSN----------------------------GKKTQALHFTSPLLTAGTESRERLTNG 535
              :|                                    |  |  |  |  |: |      :|  ||
Db        742 TNDTYDHLAKNIYQNLLFQTRLQPYLNQISFKMENDTFTLDF-SGLVQAFNHVKE--TNP 798

Qy        536 KYSYIN---KLKFGRVKNW---------QVTDGEASSKL-DFSKVIQR-----VAETEGT 577
              : ::::       | :| :::|          | : : :  |   |: ||: :        :|:| ||
Db        799 QKAFVDLAEMLAYGELRSWYEGRRLMTDYVEEAKKAGKFEDYQKVLGQETVALLAKTSGT 858

Qy        578 DEIGLIVN-----------AKAGNDDIFVGQGKMNIDGGDGHDRVFYSKDGGFGNITV- 624
                  :: |                  ||| :   |  |    ::|| |  |     : :    ||| ||
Db        859 QADDILQNVGFGHNKNVSLYGNDGNDTLIGGAGNDYLEGGSGSDTYVFGE--GFGQDTVY 916

Qy        625 ---------------DGTSA-----TEAGSY-TVNRKVARGDIYHEVVKRQETKVGKRTE 663
                             || :|      |  |:: :  |    | :  :    : :       | :
Db        917 NYDYATGRKDIIRFTDGITADMLTFTREGNHLLIKAKDGSGQVTVQSYFQNDGSGAYRID 976

Qy        664 TIQYRDYELRKVG---------------YGYQSTDNLK---SVEEVIGSQFNDVFKGSK 704
                | : : ::  |                   | |||  : |          : :  |:  :|: :  |
Db        977 EIHFDNGKVLDVATVKELVQQSTDGSDRLYAYQSGNTLNGGLGDDYLYGADGDDLLNGDA 1036

Qy        705 FNDIFHSGEGDDLLDGGAGDDRLF---------GGKGNDRLSGDEGDDLLDGGSGDDVLN 755
                ||  :|| |:| |||| |:| |:               ||:||| |:::|:| |  ||:|:| |
Db       1037 GNDSIYSGNGNDTLDGGEGNDALYGYNGNDALNGGEGNDHLNGEDGNDTLIGGAGNDYLE 1096

Qy        756 GGAGNDVYIFRKGDGNDTLYD---GTG-NDKLAFADANISDIMIERTKEG--IIVKRNDH 809
              ||:|:| |:| || | ||:|:     ||   |   : |  |     :| |:   |:||   :::|  |
Db       1097 GGSGSDTYVFGKGFGQDTVYNYDYATGRKDIIRFTDGITAD-MLTFTREGNHLLIKAKDG 1155

Qy        810 SGSINIPRWYITSNLQNYQSNKTDHKIEQLIGKDGSYITSDQIDKILQDKKDGTVITSQE 869
              || : :  ::          ||  |      ::|:::    :|   :     : :::|      ||
Db       1156 SGQVTVQYYF-----QNDGSGA--YRIDEIHFDNGKVLDVATVKELVQQSTDG------- 1201

Qy        870 LKKLADENKSQKLSASDIASSLNKLVGSMALFG 902
                         |:|  |    ::||  :|   |:|
Db       1202 ---------SDRLYAYQSGNTLNGGLGDDYLYG 1225
```

RESULT 16
S35027
cytotoxin RTX homolog frpC - Neisseria meningitidis
C;Species: Neisseria meningitidis
C;Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 08-Oct-1999
C;Accession: S35027
R;Thompson, S.A.; Wang, L.L.; Sparling, P.F.
Mol. Microbiol. 9, 85-96, 1993
A;Title: Cloning and nucleotide sequence of frpC, a second gene from Neisseria meningiti
A;Reference number: S35026; MUID:94018616; PMID:8412674
A;Accession: S35027
A;Molecule type: DNA
A;Residues: 1-1829 <THO>
A;Cross-references: GB:L06299; NID:g293961; PIDN:AAA99902.1; PID:g293963
C;Genetics:
A;Gene: frpC

```
  Query Match             9.4%;  Score 435;  DB 2;  Length 1829;
  Best Local Similarity   21.9%;  Pred. No. 3e-14;
```

```
Matches  231;  Conservative  167;  Mismatches  355;  Indels  302;  Gaps   48;

Qy         43 DFIKAAD----ELGIARLAEEPNHTETAKKSVDTVNQFL----SLTQTGIAISATKLEKF 94
Db        478 DNTKLADGSFAKHGYAALAELDSNGDNIINAADAAFQTLRVWQDLNQDGIS--------- 528

Qy         95 LQKHSTNKLAK-GLDSVE----NIDRKLGKASNVLSTLSSFLGTALAGIELDSLIKKGD- 148
Db        529 -QANELRTLEELGIQSLDLAYKDVNKNLGN-GNTLAQQGSYTKTDGTTAKMGDLLLAADN 586

Qy        149 -----------AAPDALAK--ASIDLINEI---------IGNLSQSTQTIEAFSSQLAKL 186
Db        587 LHSRFKDKVELTAEQAKAANLAGIGRLRDLREAAALSGDLANMLKAYSAAETKEAQLALL 646

Qy        187 GSTISQ-AKGFSNIGNKLQ---NLNFSKT-NLGLEIITGLLSGI--SAGFALADK----- 234
Db        647 DNLIHKWAETDSNWGKKSPMRLSTDWTQTANEGIALTPSQVAQLKKNALVSLSDKAKAAI 706

Qy        235 ----------NASTGKKVAAGFELSNQVIGNVTKAISSYV--LAQRVAAGLSTTGAVAAL 282
Db        707 DAARDRIAVLDAYTGQDSSTLYYMSEEDALNIVKVTNDTYDHLAKNIYQNL--------L 758

Qy        283 ITSSIMLAISPLAFMNAAD-----------KFNHA------NALDEFAKQFR----KFGY 321
Db        759 FQTRLQPYLNQISFKMENDTFTLDFSGLVQAFNHVKETNPQKAFVDLAEMLAYGELRSWY 818

Qy        322 DGDHLLAEY---QRGVGTIE---ASLTTISTALGAVSAGVSAAAVGSAVGTPIALLVA-- 373
Db        819 EGRRLMADYVEEAKKAGKFEDYQKVLGQETVALLAKTSGTQADDILQNVGFGHNKNVSLY 878

Qy        374 ---GVTGLISGILEASKQAMFESVANRLQGKILEWEKQNGGQNY-FDKG------YDSRY 423
Db        879 GNDGNDTLIGG---AGNDYL---------------EGGSGSDTYVFGKGFGQDTVYNYDY 920

Qy        424 AAYLANNLKFLSELNKEL-----EAERVI--------AITQQRWDNNIG----------- 459
Db        921 ATGRKDIIRFTDGITADMLTFTREGNHLLIKAKDDSGQVTVQSYFQNDGSGAYRIDEIHF 980

Qy        460 ------ELAGITKLGERIKSGKAYADAFEDGKKVEAG----------SNITLDAKTGIID 503
Db        981 DNGKVLDVATVKELVQQSTDGSDRLYAYQSGNTLNGGLGDDYLYGADGDDLLNGDAGNDS 1040

Qy        504 ISNSNGKKTQALHFTSPLLTAGTESRERLTNGKYSYINKLKFGRVKNWQVTDGEASSKLD 563
Db       1041 IYSGNGNDT----------LNGGEGNDAL------------YGYNGNDALNGG------- 1071

Qy        564 FSKVIQRVAETEGTDEIGLIVNAKAGNDDIFVGQGKMNIDGGDGHDRVFYSKDGGFGNIT 623
Db       1072 -----------EGNDHL----NGEDGNDTLIGGAGNDYLEGGSGSDTYVFGK--GFGQDT 1114

Qy        624 V----------------DGTSA-----TEAGSY-TVNRKVARGDIYHEVVKRQETKVGKR 661
Db       1115 VYNYDYATGRKDIIRFTDGITADMLTFTREGNHLLIKAKDGSGQVTVQSYFQNDGSGAYR 1174

Qy        662 TETIQYRDYELRKVG----------------YGYQSTDNLK---SVEEVIGSQFNDVFKG 702
Db       1175 IDEIHFDNGKVLDVATVKELVQQSTDGSDRLYAYQSGNTLNGGLGDDYLYGADGDDLLNG 1234

Qy        703 SKFNDIFHSGEGDDLLDGGAGDDRLF---------GGKGNDRLSGDEGDDLLDGGSGDDV 753
Db       1235 DAGNDSIYSGNGNDTLDGGEGNDALYGYNGNDALNGGEGNDHLNGEDGNDTLIGGAGNDY 1294

Qy        754 LNGGAGNDVYIFRKGDGNDTLYD---GTG-NDKLAFADANISDIMIERTKEG--IIVKRN 807
Db       1295 LEGGSGSDTYVFGKGFGQDTVYNYDYATGRKDIIRFTDGITAD-MLTFTREGNHLLIKAK 1353

Qy        808 DHSGSINIPRWYITSNLQNYQSNKTDHKIEQLIGKDGSYITSDQIDKILQDKKDGTVITS 867
Db       1354 DDSGQVTVQSYF-----QNDGSGA--YRIDEIHFDNGKVLDVATVKELVQQSTDG----- 1401

Qy        868 QELKKLADENKSQKLSASDIASSLNKLVGSMALFG 902
Db       1402 -----------SDRLYAYQSGSTLNGGLGDDYLYG 1425
```

```
RESULT 17
C81182
iron-regulated protein FrpA, probable NMB0585 [imported] - Neisseria meningitidis (strair
C;Species: Neisseria meningitidis
C;Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-Jan-2001
C;Accession: C81182
R;Tettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A
Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.;
ri, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Masignani, V.; Pizza, M.
Science 287, 1809-1815, 2000
A;Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; Ver
A;Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.
A;Reference number: A81000; MUID:20175755; PMID:10710307
A;Accession: C81182
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1302 <TET>
A;Cross-references: GB:AE002414; GB:AE002098; NID:g7225809; PIDN:AAF41013.1; PID:g722581
A;Experimental source: serogroup B, strain MC58
C;Genetics:
A;Gene: NMB0585

  Query Match             9.1%;  Score 421;  DB 2;  Length 1302;
  Best Local Similarity   22.5%;  Pred. No. 9.6e-14;
  Matches  198;  Conservative  136;  Mismatches  292;  Indels  254;  Gaps   39;

Qy        193 AKGFSNIGNKLQNLNFSKTNLGLEIITGLLSGISAGFALADKNASTGKKVAAGFELSNQV 252
Db        303 AKGFS-------GSLFDHTNNGIRTATGWVSA-DDGLLVRDLN-------------GNGI 341

Qy        253 IGNVTKAISSYVLAQRVAAGLSTTGAVAALITSSIMLAISPLAFMNAADKFNHANALDEF 312
Db        342 IDNGAELFGDNT---KLADGSFAKHGYAALAELD----------SNGDNIINAADAAFQS 388

Qy        313 AKQFRKFGYDG----DHLLAEYQRGVGTIEASLTTISTALGAVSAGVSAAAVGSAVGTPI 368
Db        389 LRVWQDLNQDGISQANELRTLEELGIQSLDLAYKDVNKNLG---NGNTLAQQGSYTKTD- 444

Qy        369 ALLVAGVTGLISGILEASKQAMFESVANRLQGKILEWEKQNGGQNYFDKG--YDSRYAAY 426
Db        445 -----GTTAKMGDLLLAA-----DNLHSRFKDKVELTAEQAKAANLAGIGRLRDLREAAA 494

Qy        427 LANNLKFLSELNKELEA-ERVIAITQ---QRW---DNNIGELA--------------GIT 465
Db        495 LSGDLANMLKAYSAAETKEAQLALLDNLIHKWAETDSNWGKKSPMRLSTDWTQTANEGIA 554

Qy        466 KLGERI------------KSGKAYADAFEDGKKVEAGSNITLDAKTG------------- 500
Db        555 LTPSQVAQLKKNALVSLSDKAKAAIDAARDRIAV-------LDAYTGQDSNTLYYMSEED 607

Qy        501 ---IIDISNSN----------------------------GKKTQALHFTSPLLTAGTES 528
Db        608 ALNIVKVTNDTYDHLAKNIYQNLLFQTRLQPYLNQISFKMENDTFTLDF-SGLVQAFNHV 666

Qy        529 RERLTNGKYSYIN---KLKFGRVKNW---------QVTDGEASSKL-DFSKVIQR----- 570
Db        667 KE--TNPQKAFVDLAEMLAYGELRSWYEGRRLMTDYVEEAKKAGKFEDYQKVLGQETVAL 724

Qy        571 VAETEGTDEIGLIVN------------AKAGNDDIFVGQGKMNIDGGDGHDRVFYSKDGG 618
Db        725 LAKTSGTQADDILQNVGFGHNKNVSLYGNDGNDTLIGGAGNDYLEGGSGSDTYVFGE--G 782

Qy        619 FGNITV----------------DGTSA-----TEAGSY-TVNRKVARGDIYHEVVKRQET 656
Db        783 FGQDTVYNYDYATGRKDIIRFTDGITADMLTFTREGNHLLIKAKDGSGQVTVQSYFQNDG 842

Qy        657 KVGKRTETIQYRDYELRKVG----------------YGYQSTDNLK---SVEEVIGSQFN 697
Db        843 SGAYRIDEIHFDNGKVLDVATVKELVQQSTDGSDRLYAYQSGNTLNGGLGDDYLYGADGD 902

Qy        698 DVFKGSKFNDIFHSGEGDDLLDGGAGDDRLF---------GGKGNDRLSGDEGDDLLDGG 748
```

```
                  |:  |   ||  :|| |:| |||| |:| |:             ||:||| |:|::|:| | ||
Db        903 DLLNGDAGNDSIYSGNGNDTLDGGEGNDALYGYNGNDALNGGEGNDHLNGEDGNDTLIGG 962

Qy        749 SGDDVLNGGAGNDVYIFRKGDGNDTLYD---GTG-NDKLAFADANISDIMIERTKEG--I 802
              :|:| | ||:|:| |:| || | | |:|:      ||  | : | |    :| |:  |:||  :
Db        963 AGNDYLEGGSGSDTYVFGKGFGQDAVYNYDYATGRKDIIRFTDGITAD-MLTFTREGNHL 1021

Qy        803 IVKRNDHSGSINIPRWYITSNLQNYQSNKTDHKIEQLIGKDGSYITSDQIDKILQDKKDG 862
              ::| | || : :  ::         || |   ::|:::  :|  :    : :::|   ||
Db       1022 LIKAKDGSGQVTVQSYF-----QNDGSGA--YRIDEIHFDNGKVLDVATVKELVQQSTDG 1074

Qy        863 TVITSQELKKLADENKSQKLSASDIASSLNKLVGSMALFG 902
                                   | :| |   ::||  :|   |:|
Db       1075 ----------------SDRLYAYQSGNTLNGGLGDDYLYG 1098
```

RESULT 18
B82736
hemolysin-type calcium binding protein XF1011 [imported] - Xylella fastidiosa (strain 9a
C;Species: Xylella fastidiosa
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000
C;Accession: B82736
R;anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen
Nature 406, 151-157, 2000
A;Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A;Reference number: A82515; MUID:20365717; PMID:10910347
A;Note: for a complete list of authors see reference number A59328 below
A;Accession: B82736
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1636 <SIM>
A;Cross-references: GB:AE003938; GB:AE003849; NID:g9105935; PIDN:AAF83821.1; GSPDB:GN001
A;Experimental source: strain 9a5c
R;Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, H
as-Neto, E.; Docena, C.; El-Dorry, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A;Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laigr
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E
A;Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;
, F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak
A;Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir
M.; Tsuhako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z
A;Reference number: A59328
A;Contents: annotation
C;Genetics:
A;Gene: XF1011

```
  Query Match             8.9%;  Score 415;  DB 2;  Length 1636;
  Best Local Similarity   22.4%;  Pred. No. 2.7e-13;
  Matches  226;  Conservative  150;  Mismatches  323;  Indels  308;  Gaps   48;

Qy         25 LYLAIPKDYDPQKGGTLNDFIKAADELG---IARLAEEPNHTETAKKSVDTVNQFLSLTQ 81
              :|||:|     |   | : | :| :|      || || |:                 | ||
Db        793 MYLALP---PRQLSATFDAFAEAYAKLKESIYIRLVLEPRLTDYLSG--------LRLTD 841

Qy         82 TG--IAISATKLEKFLQKHSTNKLAKGLDSVENIDRKLGKASNVLSTLSSFLGTALAGIE 139
                 :| |: ||  | :   :  |: |  | ::  |      || :          | :| :
Db        842 NNGVMAWDASGLEAKLDQTWQHNKAQALQDVMDLYR---YGSNAV---------AASGWK 889

Qy        140 ----LDSLIKKGDAAPD---ALAKASIDLINEIIGNLSQSTQTIEAFSSQLAKLGSTISQ 192
                  |  :| :  ||   |||  ||  ||:: |:   ||     :|          | : :
Db        890 PFDALRHMIDRTAATPDGRQALAAAEITLVS---GNAEGS----DAADLLFGDAGANLLR 942

Qy        193 AKGFSNIGNKLQNLNFSKTNLGLEIITG-----LLSGISAGFALADKNASTGKKVAAGFE 247
                 |                   | ::::|     | | :   |            |  | |
Db        943 GGG-----------------GDDVLSGGGGNDTLEGGAGNDTL---YGDDGDDVLDGGE 981

Qy        248 LSNQVIGNVTKAI--SSYVLAQRVAAGLSTTGAVAALITSSIMLAISPLAFMNAADKFNH 305
               ||:: |     :  ||  |  |  |   |     :  ||   |  :        |  |
Db        982 GSNRLEGGAGNDVLKVSYWSADNVLIG----GTGDDTLYGS---AFADTYLFNQGD--GH 1032

Qy        306 ANALDEFAKQFRKFG-----------YDGDHLLAEYQRGVGTI--------------EAS 340
               :::        ||            :|  :: :   | :|                 |
Db       1033 DTIIEQGGTDKLVFGEGILAADVRLLREGQDVVLDLGNGHDSIRLKDWLTSNGTRNHSAD 1092

Qy        341 LTTISTALGAVSAGVSAAAVG-SAVGTPIALLVAGVTG---LISG----ILEASKQAMFE 392
               :  | | | :    : :::| : :|||      :  |  |    |: |     :|: :
Db       1093 IEQIVFADGTLWTPETLSSMGLTTLGTPGNDTLKGWQGKDILLGGAGDDVLDGGEG---- 1148

Qy        393 SVANRLQG-------KILEWEKQN---GGQNYFDKGYDSRYAAYLANNLKFLSELNKELE 442
                :|||:|       |:  |   |   ||       | |: | : |:  |
Db       1149 --SNRLEGGAGDDVLKVSYWSADNVLSGG-----TGDDTLYGSAFADTYLF--------- 1192

Qy        443 AERVIAITQQRWDNNIGELAGITKL--GERIKSGKAYADAFEDGKKVEAGSNITLDAKTG 500
                      |  :  | |  |  || || | :              | : :  | :: ||
Db       1193 -------NQGDGHDTIIEQGGTDKLVFGEGILAA--------DVRLLREGQDVVLD---- 1233

Qy        501 IIDISNSNGKKTQALHFTSPLLTAGTESRERLTNGKYSYINKLKFGRVKNWQVTDGEASS 560
                                             | || |        |:|:|  ::|   :
Db       1234 -----------------------------LGNGHDSI-------RLKDWLTSNGTRNH 1255

Qy        561 KLDFSKVIQRVAETEGT-------DEIGLIVNAKAGND--------DIFV-GQGKMNIDG 604
                |   |:::   :||          :||       :|||         || : |  | :||
Db       1256 SAD----IEQIVFADGTLWTPETLSSMGLTTLGTSGNDTLKGWQGKDILLGGAGDDVLDG 1311

Qy        605 GDGHDRV------------FYSKD----GGFGNITVDGTSATEAGSYTVNRKVARGDIYH 648
              |:| :|:               ::||     || |: |: |::     | :|   |   :|| :
Db       1312 GEGSNRLEGGAGDDVLKVSYWSADNVLIGGTGDDTLYGSAF--ADTYLFN----KGDGHD 1365

Qy        649 EVVKRQET------------------------KVGKRTETIQYRDYELRKVGYGYQSTD 683
               ::::  |                        :|   ::|: :|: |     |    ||
Db       1366 TIIEQGGTDKLVFGEGILAADVRLLREGQDVVLDLGNGHDSIRLKDW-LTSNGTRNHSAD 1424

Qy        684 NLKSVEEVI-------------------GSQFNDVFKGSKFNDIFHSGEGDDLLDGGAG 723
                  |  |:::                  |    ||  ||  :   ||   |:||:|||| |
Db       1425 ----IEQIVFADGTLWTPETLSSMGLTTLGTPGNDTLKGWQGKDILLGGDGDDVLDGGEG 1480

Qy        724 DDRLFGGKGND--RLSGDEGDDLLDGGSGDDVLNGGAGNDVYIFRKGDGNDTLYDGTGND 781
               :|| || |:|  ::|    |::| ||:||| |  |  |    | |:| ||||:||: :  | |
Db       1481 SNRLEGGAGDDVLKVSYWSADNVLSGGTGDDTLYGSAFADTYLFNKGDGHDTIIEQGGTD 1540

Qy        782 KLAF-ADANISDIMIERTKEGIIVKRNDHSGSINIPRWYITSNLQNYQSNKTDHKIEQLI 840
              || | |   :  :     |: : : :  |          :|:  : |:    | : |::| |:
Db       1541 KLVFGEGAGLSRKEARFTRSGDDLSLLFNGSEDQVTVAGWF----------NGSGHQVESLV 1590

Qy        841 GKDGSYITSDQIDKILQDKKDGTVITSQELKKLADENKSQKLSASDI 887
               :||: : | ::::::          :|: :   : :   :| :| ||| |
Db       1591 FQDGT-VLSAEVERLIAAMALSPAVTTMQ-ASVGETKESPRLVASSI 1635
```

RESULT 19
C82779
hemolysin-type calcium binding protein XF0668 [imported] - Xylella fastidiosa (strain 9a
C;Species: Xylella fastidiosa
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000
C;Accession: C82779
R;anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen
Nature 406, 151-157, 2000
A;Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A;Reference number: A82515; MUID:20365717; PMID:10910347
A;Note: for a complete list of authors see reference number A59328 below
A;Accession: C82779
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1208 <SIM>
A;Cross-references: GB:AE003910; GB:AE003849; NID:g9105532; PIDN:AAF83478.1; GSPDB:GN001
A;Experimental source: strain 9a5c
R;Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, H
as-Neto, E.; Docena, C.; El-Dorry, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A;Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laigr

```
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E
A;Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;
, F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak
A;Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silvei
M.; Tsuhako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z
A;Reference number: A59328
A;Contents: annotation
C;Genetics:
A;Gene: XF0668

  Query Match             8.8%;  Score 411;  DB 2;  Length 1208;
  Best Local Similarity   22.0%;  Pred. No. 2.8e-13;
  Matches  227;  Conservative  169;  Mismatches  344;  Indels  294;  Gaps   46;

Qy         36 QKGGTLNDFIKAADELGIARLAEEPNHTETAKKSVDTVNQFL-----SLTQTG-----IA 85
Db        286 QDGQSQANELFTLSSLGIASITLTPTNTE----DVDLGNGNLIDNRGTYTRTDGRTGVVG 341

Qy         86 ISATKLEKFLQKHS-------TNKLAKGLDSVENID--RKLGKASNVLSTLSSFLGTALA 136
Db        342 DLQLGLEHFYRDYSGAHDKVTVSDAARALPAISGSGAVRDLREAASLSPALLAAVQALTP 401

Qy        137 GIELDSLIKKGDAAPDALAKASIDLINEIIGNLS------------------------- 170
Db        402 GTTRDTM------------RAALDPLVALWAGTSAMPSTEQRLETSAAVPRTVYYHGAVP 449

Qy        171 -----QSTQTIEAF-SSQLAKLGSTISQAKGFSNIGNKLQNLNFSKTNLGLEIITGLLSG 224
Db        450 AAVTAQGKEAVQAWIQQQHARLGPIIAILEKFN--GSSLVSERNGQVSTGGETFTW---- 503

Qy        225 ISAGFALADKNASTGKKVAAGFELSNQVI---GNVTKAISSYVLAQRVAAGLSTTGAVAA 281
Db        504 -------------------------NRVIHPDGHSEDVMSILLQPEQINSLMSAYASLKE 538

Qy        282 LITSSIMLA------ISPLAFMNAADKFN-HANALDEFAKQFRKFGYDGDHLLAE----Y 330
Db        539 FAYAGLVLSTRLSDYLSGLTLTYSNDAFGWDASALE--AKLDHTWQHNKTQALQDVMDLY 596

Qy        331 QRGVGTIEASLTTISTALGAVSAGVSAAAVGSAVGTPIALLVAGVTGLISGILEASKQA- 389
Db        597 RYGSNAVAASGWKPFDALRHMIDRTAATPDGRQ-----ALAAAEIT-LVSGSAEGSTAAD 650

Qy        390 --MFESVANRLQGKILEWEKQNGGQN------------YFDKGYDSRYAAYLANNLKFL 434
Db        651 LIFGDAGANILRGGAGD-DVLSGGDGDDTLYGGDGNDTLYGDAGNDTLYGG--GGNDLLL 707

Qy        435 ----------SELNKELE---AERVIAITQQRWDNNIGELAGITKLGERIKSGKAYADAF 481
Db        708 GGDGDDVLDGGEGSNRLEGGAGDDVLKVSYWSADNVL-----IGGTGDDTLYGSAFADTY 762

Qy        482 ----EDGKK--VEAGSNITLDAKTGII--DISNSNGKKTQALHFTSPLLTAGTESRERLT 533
Db        763 LFNQGDGHDTIIEQGGTDKLVFGEGIVAADVR---------------LLRQGQDVVLDLG 807

Qy        534 NGKYSYINKLKFGRVKNWQVTDGEASSKLDFSKVIQRVAETEGT-------DEIGLIVNA 586
Db        808 NGHDSI------RLKDWLTSNGTRNHSAD----IEQIVFADGTLWTPETLSSMGLTTLG 856

Qy        587 KAGND--------DIFV-GQGKMNIDGGDGHDRV-----------FYSKD----GGFGN 621
Db        857 TSGNDTLKGWQGKDILLGGAGDDVLDGGEGSNRLEGGAGNDVLKVSYWSADNVLIGGTGD 916

Qy        622 ITVDGTSATEAGSYTVNRKVARGDIYHEVVKRQET------------------------ 656
Db        917 DTLYGSAF--ADTYLFN----QGDGHDTIIEQGGTDKLVFGEGILAADVRLLREGQDVVL 970

Qy        657 KVGKRTETIQYRDYELRKVGYGYQSTDNLKSVEEVI--------------------GSQF 696
Db        971 DLGNGHDSIRLKDW-LTSNGTRNHSAD----IEQIVFADGTLWTPETLSSMGLTTLGTPG 1025

Qy        697 NDVFKGSKFNDIFHSGEGDDLLDGGAGDDRLFGGKGND--RLSGDEGDDLLDGGSGDDVL 754

Db       1026 NDTLKGWQGKDILLGGDGDDVLDGGEGSNRLEGGAGNDVLKVSYWSADNVLIGGTGDDTL 1085

Qy        755 NGGAGNDVYIFRKGDGNDTLYDGTGNDKLAFADA-NISDIMIERTKEGIIVKRNDHSGSI 813
Db       1086 YGSAFADTYLFNKGDGHDTIIEQSGTDKLVFGEGLHQKEARFTRSGDDLSLLFNGSEDQV 1145

Qy        814 NIPRWYITSNLQNYQSNKTDHKIEQLIGKDGSYITSDQIDKILQDKKDGTVITSQELKKL 873
Db       1146 TIAGWF----------NGSGHQVESLVFQDGT-VLSGEVERLIAAMALSPAATTLQ-ASV 1193

Qy        874 ADENKSQKLSASDI 887
Db       1194 GDHKESPRLVASSI 1207

RESULT 20
A47058
Fe-regulated RTX cytotoxin homolog FrpA - Neisseria meningitidis
C;Species: Neisseria meningitidis
C;Date: 21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 18-Nov-1994
C;Accession: A47058
R;Thompson, S.A.; Wang, L.L.; West, A.; Sparling, P.F.
J. Bacteriol. 175, 811-818, 1993
A;Title: Neisseria meningitidis produces iron-regulated proteins related to the RTX famil
A;Reference number: A47058; MUID:93139051; PMID:8423153
A;Contents: FAM20
A;Accession: A47058
A;Status: preliminary
A;Molecule type: nucleic acid
A;Residues: 1-1115 <THO>
A;Note: sequence extracted from NCBI backbone (NCBIN:123557, NCBIP:123558)

  Query Match             8.8%;  Score 408.5;  DB 2;  Length 1115;
  Best Local Similarity   20.4%;  Pred. No. 3.4e-13;
  Matches  239;  Conservative  171;  Mismatches  364;  Indels  399;  Gaps   51;

Qy         40 TLNDFIKAADELGIARLAEEPNHTETAKKSVDTVNQFLSL-------TQTG--------- 83
Db         55 TVNNAIQAADSF-LSSIGRDNKITNTASLLASLDNIFLNLRNVSRDIRETGKFKPNDIQQ 113

Qy         84 ----IAISA--------TKLEKFLQ-KHSTNKLAKGLDSVENIDRKLGKASNVLSTLSSF 130
Db        114 AIGDIFIAAGDGLQYIKQQTEAMAQSKFLPTKLKTGLNDV--LNSRMLKSSTVLQHELNY 171

Qy        131 LGTALA--GIEL--DSLIKKGDAAPDALAK--ASIDLINEIIGNLSQSTQTIEAFSSQLA 184
Db        172 LGFKIKDYGNERLGESIMNIDDFTPSKIANFFADPDTYSNVLEEVSRFIYSLVPDDANPW 231

Qy        185 K-----LGSTISQ------------------------------------------------- 192
Db        232 KGGEDYIGRGISEWGELLEKWYKQDFLPYLEKEWDQFPKFEDWLPEFPEWAREWLKLDPK 291

Qy        193 ------------------------AKGFSNIGNKLQNLNFSKTNLGLEIITGLLSGISAG 228
Db        292 RSGKYHVYDPLALDLDGDGIETVAAKGFA-------GALFDHRNQGIRTATGWVSA-DDG 343

Qy        229 FALADKNASTGKKVAAGFELSNQVIGNVTKAISSYVLAQRVAAGLSTTGAVAALITSSIM 288
Db        344 LLVRDLN-------------GNGIIDNGAELFGDNT---KLADGSFAKHGYAALAELD-- 385

Qy        289 LAISPLAFMNAADKFNHANALDEFAKQFRKFGYDG----DHLLAEYQRGVGTIEASLTTI 344
Db        386 --------SNGDNIINAADAAFQTLRVWQDLNQDGISQANELRTLEELGIQSLDLAYKDV 437

Qy        345 STALGAVSAGVSAAAVGSAVGTPIALLVAGVTGLISGILEASKQAMFESVANRLQGKILE 404
Db        438 NKNLG---NGNTLAQQGSYTKTD------GTTAKMGDLLLAA-----DNLHSRFKDKVEL 483

Qy        405 WEKQNGGQNYFDKG--YDSRYAAYLANNLKFLSELNKELEA-ERVIAITQ---QRW---D 455
Db        484 TAEQAKAANLAGIGRLRDLREAAALSGDLANMLKAYSAAETKEAQLALLDNLIHKWAETD 543

Qy        456 NNIGELA--------------GITKLGERI------------KSGKAYADAFEDGKKVEA 489
```

```
Db     544 SNWGKKSPMRLSTDWTQTANEGIALTPSQVAQLKKNALVSLSDKAKAAIDAARDRIAV-- 601

Qy     490 GSNITLDAKTG---------------IIDISNSN------------------------ 508

Db     602 -----LDAYTGQDSSTLYYMSEEDALNIVKVTNDTYDHLAKNIYQNLLFQTRLQPYLNQI 656

Qy     509 ----GKKTQALHFTSPLLTAGTESRERLTNGKYSYIN---KLKFGRVKNW---------Q 552

Db     657 SFKMENDTFTLDF-SGLVQAFNHVKE--TNPQKAFVDLAEMLAYGELRSWYEGRRLMADY 713

Qy     553 VTDGEASSKL-DFSKVIQR-----VAETEGTDEIGLIVN------------AKAGNDDIF 594

Db     714 VEEAKKAGKFEDYQKVLGQETVALLAKTSGTQADDILQNVGFGHNKNVSLYGNDGNDTLI 773

Qy     595 VGQGKMNIDGGDGHDRVFYSKDGGFGNITV----------------DGTSA-----TEAG 633

Db     774 GGAGNDYLEGGSGSDTYVFGK--GFGQDTVYNYDYATGRKDIIRFTDGITADMLTFTREG 831

Qy     634 SY-TVNRKVARGDIYHEVVKRQETKVGKRTETIQYRDYELRKVG--------------- 676

Db     832 NHLLIKAKDDSGQVTVQSYFQNDGSGAYRIDEIHFDNGKVLDVATVKELVQQSTDGSDRL 891

Qy     677 YGYQSTDNLK---SVEEVIGSQFNDVFKGSKFNDIFHSGEGDDLLDGGAGDDRLF----- 728

Db     892 YAYQSGSTLNGGLGDDYLYGADGNDLLNGDAGNDSIYSGNGNDTLDGGEGNDALYGYNGN 951

Qy     729 ----GGKGNDRLSGDEGDDLLDGGSGDDVLNGGAGNDVYIFRKGDGNDTLYD---GTGND 781

Db     952 DALNGGEGNDHLNGEDGNDTLIGGAGNDYLEGGSGSDTYVFGEGFGQDTVYNYHVDKNSD 1011

Qy     782 KLAFADANISDIMIERTKEGIIVKRNDHSGSINIPRWYITSNLQNYQSNKTDHKIEQLIG 841

Db    1012 TMHFKGFKAADVHFIRSGSDLVLSASEQD-NVRISGFFYGEN----------HRVDTFVF 1060

Qy     842 KDGSYITSDQIDKILQDKKDGTVITSQELKKLADENKSQKLSASDIASSL---NKLVGSM 898

Db    1061 DDAA----------------------------------ISNPDFAKYINAGNNLVQSM 1084

Qy     899 ALFG-----TANSVSSNALQPITQPTQGILAPS 926

Db    1085 SVFGSNTAATGGNVDAN-IQSVQQPL--LVTPS 1114
```

RESULT 21
G82562
bacteriocin XF2407 [imported] - Xylella fastidiosa (strain 9a5c)
C;Species: Xylella fastidiosa
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000
C;Accession: G82562
R;anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen
Nature 406, 151-157, 2000
A;Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A;Reference number: A82515; MUID:20365717; PMID:10910347
A;Note: for a complete list of authors see reference number A59328 below
A;Accession: G82562
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-2064 <SIM>
A;Cross-references: GB:AE004049; GB:AE003849; NID:g9107579; PIDN:AAF85206.1; GSPDB:GN001
A;Experimental source: strain 9a5c
R;Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, H
as-Neto, E.; Docena, C.; El-Dorry, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A;Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laigr
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E
A;Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;
, F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak
A;Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir
M.; Tsuhako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z

A;Reference number: A59328
A;Contents: annotation
C;Genetics:
A;Gene: XF2407

```
  Query Match             8.4%;  Score 391;  DB 2;  Length 2064;
  Best Local Similarity   25.4%;  Pred. No. 6.5e-12;
  Matches  135;  Conservative   89;  Mismatches  163;  Indels  144;  Gaps   23;

Qy     446 VIAITQQRWDNNIGELAGITKLGERIKSGKAYADAF----EDGKK--VEAGSNITLDAKT 499

Db    1588 VLKVAYRSADNVL-----IGGTGDDTLYGSAYADTYLFNKGDGHDTIIEQGGDDTLVFGA 1642

Qy     500 GIIDISNSNGKKTQALHFTSPLLTAGTESRERLTNGKYSYINKLKFGRVKNWQVTDGEAS 559

Db    1643 GIV--------ASQV-----RVLREGQDVVLDLGNGHDSI-------RLKDWLTSDGYRN 1682

Qy     560 SKLDFSKVIQRVAETEGT-------DEIGLIVNAKAGNDDIFVGQGKMNIDGGDGHDRVF 612

Db    1683 GNND----IEQIVFADGTIWTPETLSSMGLTTLGTSGNDTLKGWQGKDILLGGDGDDVL- 1737

Qy     613 YSKDGGFG--------------------NITVDGT------SATEAGSYTVNRKVARGD 645

Db    1738 ---DGGMGSNRLEGGAGNDVLKVAYRSADNVLIGGTGDDTLYGSGYADTYLFN----KGD 1790

Qy     646 IYHEVVKR--QETKV----------------------GKRTETIQYRDYELRKVGYGYQ 680

Db    1791 GHDTIIEQGGDDTLVFGAGIVASQVRVLREGQDVVLDLGNGHDSIRLKDW---LTSDGYR 1847

Qy     681 STDNLKSVEEVI--------------------GSQFNDVFKGSKFNDIFHSGEGDDLLDG 720

Db    1848 NGNN--DIEQIVFADGTIWTPETLSSMGLTTLGTSGNDTLKGWQGKDILLGGDGDDVLDG 1905

Qy     721 GAGDDRLFGGKGND--RLSGDEGDDLLDGGSGDDVLNGGAGNDVYIFRKGDGNDTLYDGT 778

Db    1906 GMGSNRLEGGAGNDVLKVAYRSADNVLIGGTGDDTLYGSAYADTYLFNKGDGHDTIIEQG 1965

Qy     779 GNDKLAFADANISDIMIERTKEG--IIVKRNDHSGSINIPRWYITSNLQNYQSNKTDHKI 836

Db    1966 GDDTLVFG-AGLHQKEARFTKSGNDLSILFNASEDQVTIAGWF----------NGSGHQV 2014

Qy     837 EQLIGKDGSYITSDQIDKILQDKKDGTVITSQELKKLADENKSQKLSASDI 887

Db    2015 ESLVFQDGT-VLSGEVERLIAAMALSSAVTTTQ-ASVRDTKESHRLVASSI 2063
```

RESULT 22
S34238
leukotoxin A - Pasteurella haemolytica (fragment)
N;Alternate names: lktA protein
C;Species: Pasteurella haemolytica
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 23-Mar-2001
C;Accession: S34238; S34236
R;Lainson, A.F.; Aitchison, K.D.; Donachie, W.
submitted to the EMBL Data Library, June 1993
A;Description: DNA sequence of the carboxy terminal end of leukotoxin A from the T4 serc
of Pasturella haemolytica.
A;Reference number: S34236
A;Accession: S34238
A;Molecule type: DNA
A;Residues: 1-208 <LAI>
A;Cross-references: EMBL:Z22885; NID:g311830; PIDN:CAA80499.1; PID:g311831
A;Experimental source: serotype T4
A;Accession: S34236
A;Molecule type: DNA
A;Residues: 9-208 <LA2>
A;Cross-references: EMBL:Z22886; NID:g311826; PIDN:CAA80500.1; PID:g311827
A;Experimental source: serotype T15
C;Genetics:
A;Gene: lktA
C;Superfamily: hemolysin A; hemolysin A homology
C;Keywords: calcium binding; cytolysis; exotoxin; hemolysis; lipoprotein; tandem repeat

```
  Query Match             8.2%;  Score 383;  DB 2;  Length 208;
  Best Local Similarity   37.8%;  Pred. No. 6.1e-13;
  Matches   79;  Conservative   52;  Mismatches   64;  Indels   14;  Gaps    8;

Qy        723 GDDRLFGGKGNDRLSGDEGDDLLDGGSGDDVLNGGAGNDVYIFRKGDGNDTLYDGTGNDK 782
              |:|||||||:|  :  |   ||| :|||  |:|:|:||  |:|:::  |:|||||::  :   ||||
Db          1 GNDRLFGGKGDDIIDGGNGDDFIDGGKGNDLLHGGKGDDIFVHRQGDGNDSITESEGNDK 60

Qy        783 LAFADANISDIMIERTKEGIIVKRNDHSGSINIPRWY----ITSNLQNYQSNKTDHKIEQ 838
              |:|:|:|:  |:     :::  |        |    |:         ::||  :  :  |  |||:
Db         61 LSFSDSNLKDLTFEKVNHHLVI-TNTKQEKVTIQNWFREAEFAKTIRNYVATR-DDKIEE 118

Qy        839 LIGKDGSYITSDQIDKILQDKKDGTVITSQELKKLADENKSQKLSASDIASSLNKLVGSM 898
              :||::|   |||  |:|::::  |      |     :|  ::  |   :   |   |   | ::||:||: |
Db        119 IIGQNGERITSKQVDELIEKGKG--KIDKSDLSQVVDNYQLLKYS-RDASNSLDKLISSA 175

Qy        899 ALFGTANSVSSNALQPITQPTQGILAPSV 927
              :  |  ::|  | | |      ||  :|  ||:
Db        176 SAFTSSND-SRNVL---ASPT-SMLDPSL 199
```

RESULT 23
F81856
probable RTX-family exoprotein NMA1626 [imported] - Neisseria meningitidis (strain Z2491
C;Species: Neisseria meningitidis
C;Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 02-Feb-2001
C;Accession: F81856
R;Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel
; Holroyd, S.; Jagels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandream,
Nature 404, 502-506, 2000
A;Title: Complete DNA sequence of a serogroup A strain of Neisseria menigitidis Z2491.
A;Reference number: A81775; MUID:20222556; PMID:10761919
A;Accession: F81856
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-697 <PAR>
A;Cross-references: GB:AL162756; GB:AL157959; NID:g7380091; PIDN:CAB84854.1; PID:g738026
A;Experimental source: serogroup A, strain Z2491
C;Genetics:
A;Gene: NMA1626

```
  Query Match             7.9%;  Score 366;  DB 2;  Length 697;
  Best Local Similarity   21.1%;  Pred. No. 2.6e-11;
  Matches  201;  Conservative  130;  Mismatches  294;  Indels  328;  Gaps   37;

Qy         45 IKAADELGIARLAEEPNHTETAKKSVDTVNQFLSLTQTG--------IAISATKLEKFLQ 96
              ::   :||||  |              : :|:  |   | |:|            : ::|   |
Db          1 MRTLEELGIQSLDLTYKDVNKNLGNGNTLAQQGSYTKTNGTTAKMGDLLLAADNL----- 55

Qy         97 KHSTNKLAKGLDSVENIDRKLGKASNVLSTLSSFLGTALAGIELDSLIKKGDAAPDALAK 156
               ||  |       | ||  :  :  ||:|               ||||        :::  |    ||
Db         56 -HSRFK-----DKVE-LTAEQAKAAN------------LAGIGRLRDLREAAALSGDLA- 95

Qy        157 ASIDLINEIIGNLSQSTQTIEAFSSQLAKLGSTISQ-AKGFSNIGNKLQNLNFSKTNLGL 215
                          |: ::       :|||  |  :  |  |:   ||  ||                   :
Db         96 -----------NMLKAYSAAETKEAQLALLDNLIHKWAETDSNWGKKSP----------M 134

Qy        216 EIITGLLSGISAGFALADKNASTGKKVAAGFELSNQVIGNVTKAISSYVLAQ-RVAAGLS 274
               :  |        : | ||        :   ||            | ::       ||  ::     |: |:|      :
Db        135 RLSTDWTQTANEGIALTPSQVAQLKK--------NALVSLSDKAKAAIDAARDRIAVLDA 186

Qy        275 TTGAVAALITSSIMLAISPLAFMNAADKFNHANALDEFAKQFRKFGYDGDHLLAEYQRGV 334
               ||   :              | |  :|:   |    |              ::           |     |      :||  :
Db        187 YTGQDS-----------STLYYMSEEDALNIVKVTNDTYDHLAKNIY--QNLLFQ----- 228

Qy        335 GTIEASLTTISTALGAVSAGVSAAAVGSAVGTPIALLVAGVTGLISGILEASKQAMFESV 394
                       | :    |   :|   :                          |  :|:      :  :  |  :  |    |     :
Db        229 -------TRLQPYLNQISFKME--------NDTFTLDFSGLVQAFNHVKETNPQKAFVDL 273

Qy        395 ANRLQGKILEWEKQNGGQNYFDKGYDSRYAAYLANNLKFLSELNKELEAERVIAITQQRW 454
               |    |                                      ||           ||       |    | ::|
Db        274 AEML--------------------------AY--------GELRSWYEGRRLMA------ 293

Qy        455 DNNIGELAGITKLGERIKSGKAYADAFEDGKKVEAGSNITLDAKTGIIDISNSNGKKTQA 514
                                :  |   |:||     ||| :||        :  |  ||
Db        294 ---------DYVEEAKKAGK-----FEDYQKVLGQETVALLAK--------------- 322

Qy        515 LHFTSPLLTAGTESRERLTNGKYSYINKLKFGRVKNWQVTDGEASSKLDFSKVIQRVAET 574
                          |:||::   :  |   |                  :   ||    ||
Db        323 --------TSGTQADDILQN--------VGFGHNKN--------------------VSL 345

Qy        575 EGTDEIGLIVNAKAGNDDIFVGQGKMNIDGGDGHDRVFYSKDGGFGNITV---------- 624
               |  |                   |||  :     |  |      ::||  |   |          :   |   |||     ||
Db        346 YGND----------GNDTLIGGAGNDYLEGGSGSDTYVFGK--GFGQDTVYNYDYATGRK 393

Qy        625 ------DGTSA-----TEAGSY-TVNRKVARGDIYHEVVKRQETKVGKRTETIQYRDYEL 672
                            || :|        |   |:   :     |       |    :    : :         |   :   |   :  :  ::
Db        394 DIIRFTDGITADMLTFTREGNHLLIKAKDDSGQVTVQSYFQNDGSGAYRIDEIHFDNGKV 453

Qy        673 RKVG----------------YGYQSTDNLK---SVEEVIGSQFNDVFKGSKFNDIFHSGE 713
               |                                 |  |||  :  |           :   :  |:    :|:      |         ||       :||
Db        454 LDVATVKELVQQSTDGSDRLYAYQSGNTLNGGLGDDYLYGADGDDLLNGDAGNDSIYSGN 513

Qy        714 GDDLLDGGAGDDRLF---------GGKGNDRLSGDEGDDLLDGGSGDDVLNGGAGNDVYI 764
               |:|   |:||  |:|   |:                      ||:|||   |:|::|:|   |   ||:|:|   |   |||:|:|   |:
Db        514 GNDTLNGGEGNDALYGYNGNDALNGGEGNDHLNGEDGNDTLIGGAGNDYLEGGSGSDTYV 573

Qy        765 FRKGDGNDTLYD---GTGNDKLAFADANISDIMIERTKEGIIVKRNDHSGSINIPRWYIT 821
               |  ||  |     ||:|:                                 :|:        :    :::       ::         :   ::
Db        574 FGKGFGQDTVYNYHYVDKNSDTMHFKGFKAADVHFIRSGSDLVLSASEQD-NVRISGFFYG 632

Qy        822 SNLQNYQSNKTDHKIEQLIGKDGSYITSDQIDKILQDKKDGTVITSQELKKLADENKSQK 881
               |                       |:::    :     |    :
Db        633 EN----------HRVDTFVFDDAA-------------------------------------- 646

Qy        882 LSASDIASSL---NKLVGSMALFG-----TANSVSSNALQPITQPTQGILAPS 926
              :|   |  |     :      | ||  ||::||      |    :|    :|   :|    :    ||        :: ||
Db        647 ISNPDFAKYINAGNNLVQSMSVFGSNTAATGGNVDAN-IQSVQQPL--LVTPS 696
```

RESULT 24
AH2515
hypothetical protein alr7304 [imported] - Nostoc sp. (strain PCC 7120) plasmid pCC7120al
C;Species: Nostoc sp. PCC 7120
A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002
C;Accession: AH2515
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi,
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S
DNA Res. 8, 205-213, 2001
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anal
A;Reference number: AB1807; MUID:21595285; PMID:11759840
A;Accession: AH2515
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-4936 <KUR>
A;Cross-references: GB:BA000020; PIDN:BAB78388.1; PID:g17135842; GSPDB:GN00180
A;Experimental source: strain PCC 7120
C;Genetics:
A;Gene: alr7304
A;Genome: plasmid

```
  Query Match             7.4%;  Score 344;  DB 2;  Length 4936;
  Best Local Similarity   22.1%;  Pred. No. 6e-09;
  Matches  233;  Conservative  133;  Mismatches  360;  Indels  330;  Gaps   48;

Qy         24 NLYLAIPKDYDPQKGGTLNDFIKA-------ADELGIA----RLAEEPNHTETAKKSVDT 72
              |  :  :  |  :           |:   |   ||                 ::     ||              :   :   |  |  ||
Db       3662 NSDIIVLKVFKDNGAGSFADLEKALQWVAANSNTYNIASVNLSIGDSQNWT-TATGRYGI 3720

Qy         73 VNQFLSLTQTGIAISATKLEKFLQKHST----------NKLAKGLDSVENI---DRKLGK 119
               ::   ::        |  |:|         |  |   |                     |:|  |        :|               :|
Db       3721 GDELSAIASNNIIINAAAGNSFYQYTSNPGLAYPAIDPNVIAVGAVWADNFGGPKNFVGG 3780

Qy        120 ASNVLSTLSSFLGTALAGIELDSLIKKGDAAPDALAKASIDLINEIIGNLSQSTQTIEAF 179
               | :  :|          :       ||    :       |           ||                      :|   ||:|          |:
```

```
Db       3781 AIDYTTTADQIASFSQRHPELLDIFAPGILITGANANGG----TTTLGGTSQAT----AY 3832

Qy        180 SSQLAKLGSTISQAKGFSNIGNKL-------------------QNLNFSKTNLG------ 214
Db       3833 LTGVATLAQQIAQEK----LGRKLTVTEFRNLLDTTSVIINDGDNENDNVTNTGFNYPRV 3888

Qy        215 --LEIITGLLS---------GISAGFALADKNASTGK------------KVAAGFELSN 250
Db       3889 DLLKLAEAILSLTGTTPNPDPVNPG----NNNNNNGTTTSDNTINQVHTVNLAAGQVRTD 3944

Qy        251 QVIGNVTKAISSYVLAQRVAAGLSTTGAVAALITSSIMLAISPLAFMNAADKFNHANALD 310
Db       3945 VDFGN-----------QQIITNQAPT--VANAIADQIINEDANFTFVIPANTFVDADAGD 3991

Qy        311 ------------EFAKQFRKF-GYDGDHLLAEYQRGVGTIEASLTTISTALGAVSAGVSA 357
Db       3992 VLTYSTTLPSWLTFNATTRTFSGTPGN-------SNVGTVNITVT-------------AT 4031

Qy        358 AAVGSAVGTPIALLVAGVTGL-ISGILEASKQA------MFESVANRLQGKILEWEKQNG 410
Db       4032 DSTGASVDDSFTLTVANTNDAPILGLAIADQSTASNTPFTFQIPLNTFSD---------- 4081

Qy        411 GQNYFDKGYDSRYAAYLANNLKFLSELNKELEAERVIAITQQRWDNNIGELAGITKLGER 470
Db       4082 ----IDTGDTLTYSAKLVGDIPLPTWL------------TFNATNRTFSGIPG------- 4118

Qy        471 IKSGKAYADAFEDGKKVEAGS-NITLDAKTGIIDISNSNGKKTQALHFTSPL-------- 521
Db       4119 ---------------NVDVGTLNITVQA----IDTSNASISDSFVLTITNLINNIVGTSG 4159

Qy        522 --LTAGTESRERL--TNGKYSYINKLKFGRVKNWQVTDGEASSKL-------------DF 564
Db       4160 NNTLAGTPNNDNIQGLGG-----NDIIFGLAGNDTLNGGTGSDTMTGGLGDDTYIVDNNV 4214

Qy        565 SKVIQRVAETEGTDEI-------------GLI-------------------------VN 585
Db       4215 DKVVENL--NEGIDTVRSSISYTLLENVENLILTGTSNISGTGNILSNIITGNSGANTLN 4272

Qy        586 AKAGNDDIFVGQ-GKMNIDGGDGHDRVFYSKDGGFGNITVDGTSATEAGSYTVNRKVARG 644
Db       4273 GKAG-DDILNGEGGNDNLKGEDGNDVL----NGGAGNDILDGGLGDDVMTGGVG-----N 4322

Qy        645 DIYH----------EVVKRQETKVGKRTETIQYRDYELRKVG-YGYQSTDN-LKSVEEVI 692
Db       4323 DIYYVDSSNDIIIDELNEGTDTVNTIITWTLGNHLENLTLIGSSAINGTGNALKNI--II 4380

Qy        693 GSQFNDVFKGSKFNDIFHSGEGDDLLDGGAGDDRLFGGKGNDRLSGDEGDDLLDGGSGDD 752
Db       4381 GNSADNILSGGDNDDILRGGEGNDTLYGGAGNDSLDGGIGNDSLNGEDGNDNLKGDVGND 4440

Qy        753 VLNGGAGNDVYIFRKGDGNDTLYDGTGNDKLAFADANISDIMIERTKEGIIVKRNDHSGS 812
Db       4441 ILNGNAGNDT--LDGGLGDDVMTGGAGND-IYFVDSS-NDTIIEELNEG----TDTVNAS 4492

Qy        813 INIPRWYITSNLQNYQSNKTDHKIEQLIGKDGSYITSDQIDKILQDKKDGTVITSQELKK 872
Db       4493 IN---WTLGNNLENL----------TLTGSNGINGTGNALKNI----------------- 4522

Qy        873 LADENKSQKLSASDIASSLNKLVGSMALFGTANSVS 908
Db       4523 ITGNNGDNILSGGDNDDTLRGNAGNDTLFGGSGNDS 4558
```

RESULT 25
AH3098
rhizobiocin/RTX toxin [imported] - Agrobacterium tumefaciens (strain C58, Dupont)
C;Species: Agrobacterium tumefaciens
C;Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 18-Nov-2002
C;Accession: AH3098
R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, L
erage, G.; Gillet, W.; Grant, C.; Guenthner, D.; Kutyavin, T.; Levy, R.; Li, M.; McClell
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
ster, E.W.
A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A;Reference number: AB2577; MUID:21608550; PMID:11743193
A;Accession: AH3098
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1944 <KUR>
A;Cross-references: GB:AE008689; PIDN:AAL45206.1; PID:g17742885; GSPDB:GN00187
A;Experimental source: strain C58 (Dupont)
C;Genetics:
A;Gene: rzcA
A;Map position: linear chromosome

```
  Query Match             7.1%;  Score 330.5;  DB 2;  Length 1944;
  Best Local Similarity   21.7%;  Pred. No. 7.7e-09;
  Matches  218;  Conservative  147;  Mismatches  333;  Indels  305;  Gaps   48;

Qy         66 AKKSVDTVNQFLSLTQTG-----IAISATK------LEKFLQKHST-NKLAKGLDSVENI 113
Db        653 ARAMADAGDDIVASAASGASVANVALSAAKSLRSQDIRSFVWTDSNGDVVAEGKEAI--A 710

Qy        114 DRKLGKASNVLSTLSSFLGTALAGIE----LDSLIKKGDAAPDALAKASIDLINEIIGNL 169
Db        711 DLSLGKHKLTLTT------TDVNGVKTSDTLDVLV-RGD--------------------- 742

Qy        170 SQSTQTIEAFSSQLAKLGSTISQAKGF----SNIGN-KLQNLNFSKTNLGLEIITGLLSG 224
Db        743 --RTLLVDNFND-----GKADGWATDFGATDGNTGNFLLKGTVFSRSTTG--------DG 787

Qy        225 ISA-GFALADKNASTGKKVAAGFELSNQVIGNVTKAISSYVLAQRVAAGLSTTGAVAALI 283
Db        788 LDAPEAALFDQSDATGNKLV--------YIGAQSSGWSNYVF-------------EATL 825

Qy        284 TSSIMLAISPLAFMNAADKFNHANALDEFAKQFRKFGYDGD---HLLAEYQRGVGTIEAS 340
Db        826 TQLDNDAIG--VYFYYSD-----------AKNYYRFTMDGEANRRQLVKVSNGVETLLAS 872

Qy        341 LT-----------TISTALGAVSAGVSAAAVGSAVGTPIALLVAGVTGLISGILEASKQA 389
Db        873 VNEGSPYNMDIPLTVAIVDGAINVFLGDKNV---FGGPVVDVTSPLSGGTVGLYSSGQRA 929

Qy        390 -MFESVANRLQGKILEWEKQNGGQNYFDKGYDSRYAAYLANNLKFLSELNKELEAERVIA 448
Db        930 SVFDDI---VVTKAATTAKAGIDQRVYDFDGDGK--------------VSVRLDASSSFS 972

Qy        449 ITQQR---WDNNIGELAGITKLGERIKSGKAYADAFEDGKKVEAGSN------------- 492
Db        973 AGQLRDFVWTDLDGNVVAI--------------------GKKVDADLNVGVNKLLLMVSDA 1013

Qy        493 ----------ITLDAKTGIIDISN----------------------SNGKKTQALHFTSP 520
Db       1014 QGSVSTDRIDVTVVDKTKILVAENFSTAEAMARFKIVDEGEFGGIGADGKSSEWLISDGK 1073

Qy        521 LL-TAGTESRERLTNG-----KYSY--------INKLKFGRVKNWQVTDGEASSKLDFSK 566
Db       1074 LLQTTGLMSRELTWNGATNSDKYKRGWSPMGDGVNVLRLGTYALF--NDPAAQAWTDYA- 1130

Qy        567 VIQRVAETEGTDEIGLIVNAKAGNDDIFVGQGKMNIDGGDGHDRVFYSKDGGFGNI---- 622
Db       1131 -IEANFQTPDKDGLGFLFRYKDSKN-----YYKLELDADGILDR---NPGNGAGSIFNLV 1181

Qy        623 ----TVDGTSATEAGSYTVNRKVARGDIYHEVVKRQETKVGKRTETIQY--RDYELRKVG 676
Db       1182 RMKNGIEEILAQVPGKYEPGQEM---KLRVEVIGDKITAFLNEEALFAYPIGDRGLAAGT 1238

Qy        677 YGYQSTDN------------LKS---VEEVIGSQFNDVFKGS---------KFNDIFHSG 712
Db       1239 FGLYSWGNAGLTFDNLTVVDLKSGLEVNRIVGTNGADVLVGTAAAETMFGLEGNDELQGF 1298

Qy        713 EGDDLLDGGAGDDRLFGGKGNDRLS---------GDEGDDLLDGGSGDDVLNGGAGNDV- 762
Db       1299 GGDDRLDGGAGDDALAGGTGNDALSGGAGNDELWGDDGDDMIAGGLGDDFIEGGRGNDLL 1358
```

```
Qy        763 --------YIFRKGDGNDTLYD---GTGN-DKLAFADANISDIMIERTKEGIIVKRNDHS 810
                      | : :|||:|  : :     :|: |:|:  | : |: :: :   : :::: |
Db       1359 IGGDGSDRYRYGRGDGSDVIVETASASGSVDRLSLYDIDRSEAVLRKYGQSVVIELADGE 1418

Qy        811 GSINIPRWYITSNLQNYQSNKTDHKIEQLIGKDGSYITSDQIDKILQDKKDGTVITSQEL 870
                        | | |      |  ||:|  ||   :     | | | ::  | |     |
Db       1419 ----------TLTLSN---QLADGGIERLSFADGVVLNRGDIVKGLVNR--GPVAADDGL 1463

Qy        871 KKLADENKSQKLS-ASDIASSLNKLVGSMALFGTANSVSSNAL 912
                 : ::  |   :| |: : :  :  :  : : | :  |   |:
Db       1464 AAVNEDAPSFVISFATLLGNDRDADLDGLTVTGVSAPVGGTAV 1506
```

RESULT 26
A96188
probable phosphoesterase (EC 3.1.-.-) yvnB [imported] - Agrobacterium tumefaciens (strai
C;Species: Agrobacterium tumefaciens
C;Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 18-Nov-2002
C;Accession: A96188
R;Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman,
A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.;
Science 294, 2323-2328, 2001
A;Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum
A;Reference number: A97359; MUID:21608551; PMID:11743194
A;Accession: A96188
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1990 <KUR>
A;Cross-references: GB:AE007870; PIDN:AAK89027.1; PID:g15158819; GSPDB:GN00170
C;Genetics:
A;Gene: AGR_L_909
A;Map position: linear chromosome

```
  Query Match             7.1%;  Score 330.5;  DB 2;  Length 1990;
  Best Local Similarity   21.7%;  Pred. No. 8e-09;
  Matches  218;  Conservative  147;  Mismatches  333;  Indels  305;  Gaps   48;

Qy         66 AKKSVDTVNQFLSLTQTG-----IAISATK------LEKFLQKHST-NKLAKGLDSVENI 113
              |:   |  :  ::   :|       :|:|| |       :   |:    |  : :|:|  :::
Db        699 ARAMADAGDDIVASAASGASVANVALSAAKSLRSQDIRSFVWTDSNGDVVAEGKEAI--A 756

Qy        114 DRKLGKASNVLSTLSSFLGTALAGIE----LDSLIKKGDAAPDALAKASIDLINEIIGNL 169
              |  |||     |:|       | : |::      || |: :||
Db        757 DLSLGKHKLTLTT------TDVNGVKTSDTLDVLV-RGD--------------------- 788

Qy        170 SQSTQTIEAFSSQLAKLGSTISQAKGF----SNIGN-KLQNLNFSKTNLGLEITGLLSG 224
                 | :: |:        |         |      | ||  |:     ||:: |           |
Db        789 --RTLLVDNFND-----GKADGWATDFGATDGNTGNFLLKGTVFSRSTTG--------DG 833

Qy        225 ISA-GFALADKNASTGKKVAAGFELSNQVIGNVTKAISSYVLAQRVAAGLSTTGAVAALI 283
              : |    || |:: :|| |:                  ||    | :||              |  :
Db        834 LDAPEAALFDQSDATGNKLV--------YIGAQSSGWSNYVF--------------EATL 871

Qy        284 TSSIMLAISPLAFMNAADKFNHANALDEFAKQFRKFGYDGD---HLLAEYQRGVGTIEAS 340
              |      ||     :     :|             || : :|  ||:      |  :    || |: ||
Db        872 TQLDNDAIG--VYFYYSD-----------AKNYYRFTMDGEANRRQLVKVSNGVETLLAS 918

Qy        341 LT-----------TISTALGAVSAGVSAAAVGSAVGTPIALLVAGVTGLISGILEASKQA 389
              :               |::    ||:: :      |       | |:  : : ::|     |:  : ::|
Db        919 VNEGSPYNMDIPLTVAIVDGAINVFLGDKNV---FGGPVVDVTSPLSGGTVGLYSSGQRA 975

Qy        390 -MFESVANRLQGKILEWEKQNGGQNYFDKGYDSRYAAYLANNLKFLSELNKELEAERVIA 448
               :|: :   :    :     |     :|  |   :           ::  |:|   :
Db        976 SVFDDI---VVTKAATTAKAGIDQRVYDFDGDGK--------------VSVRLDASSSFS 1018

Qy        449 ITQQR---WDNNIGELAGITKLGERIKSGKAYADAFEDGKKVEAGSN------------- 492
                || |   |   :   |                            ||||:|  |
Db       1019 AGQLRDFVWTDLDGNVVAI-------------------GKKVDADLNVGVNKLLLMVSDA 1059

Qy        493 ----------ITLDAKTGIIDISN----------------------SNGKKTQALHFTSP 520
                        :|:  || |:  |                        ::|| :: |
Db       1060 QGSVSTDRIDVTVVDKTKILVAENFSTAEAMARFKIVDEGEFGGIGADGKSSEWLISDGK 1119

Qy        521 LL-TAGTESRERLTNG-----KYSY--------INKLKFGRVKNWQVTDGEASSKLDFSK 566
              || |  |  |||   ||        ||              :| |: |      :   |  |  : |::
Db       1120 LLQTTGLMSRELTWNGATNSDKYKRGWSPMGDGVNVLRLGTYALF--NDPAAQAWTDYA- 1176

Qy        567 VIQRVAETEGTDEIGLIVNAKAGNDDIFVGQGKMNIDGGDGHDRVFYSKDGGFGNI---- 622
               |:  :|  |  :|: :  |  :        |: :|   ||   :   | |:|
Db       1177 -IEANFQTPDKDGLGFLFRYKDSKN-----YYKLELDADGILDR---NPGNGAGSIFNLV 1227

Qy        623 ----TVDGTSATEAGSYTVNRKVARGDIYHEVVKRQETKVGKRTETIQY--RDYELRKVG 676
                   ::    |   | |    :::    :  ||:  :  |                |    |   |
Db       1228 RMKNGIEEILAQVPGKYEPGQEM---KLRVEVIGDKITAFLNEEALFAYPIGDRGLAAGT 1284

Qy        677 YGYQSTDN------------LKS---VEEVIGSQFNDVFKGS---------KFNDIFHSG 712
              :|  |  |          |||   |  ::|:   ||  |:              :  ||
Db       1285 FGLYSWGNAGLTFDNLTVVDLKSGLEVNRIVGTNGADVLVGTAAAETMFGLEGNDELQGF 1344

Qy        713 EGDDLLDGGAGDDRLFGGKGNDRLS---------GDEGDDLLDGGSGDDVLNGGAGNDV- 762
                 |||  ||||||||  | || |||  ||               ||:|||:: || |||  :  || |||:
Db       1345 GGDDRLDGGAGDDALAGGTGNDALSGGAGNDELWGDDGDDMIAGGLGDDFIEGGRGNDLL 1404

Qy        763 --------YIFRKGDGNDTLYD---GTGN-DKLAFADANISDIMIERTKEGIIVKRNDHS 810
                      | : :|||:|  : :     :|: |:|:  | : |: :: :   : :::: |
Db       1405 IGGDGSDRYRYGRGDGSDVIVETASASGSVDRLSLYDIDRSEAVLRKYGQSVVIELADGE 1464

Qy        811 GSINIPRWYITSNLQNYQSNKTDHKIEQLIGKDGSYITSDQIDKILQDKKDGTVITSQEL 870
                        | | |      |  ||:|  ||   :     | | | ::  | |     |
Db       1465 ----------TLTLSN---QLADGGIERLSFADGVVLNRGDIVKGLVNR--GPVAADDGL 1509

Qy        871 KKLADENKSQKLS-ASDIASSLNKLVGSMALFGTANSVSSNAL 912
                 : ::  |   :| |: : :  :  :  : : | :  |   |:
Db       1510 AAVNEDAPSFVISFATLLGNDRDADLDGLTVTGVSAPVGGTAV 1552
```

RESULT 27
C82521
hemolysin-type calcium binding protein XF2759 [imported] - Xylella fastidiosa (strain 9a
C;Species: Xylella fastidiosa
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000
C;Accession: C82521
R;anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen
Nature 406, 151-157, 2000
A;Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A;Reference number: A82515; MUID:20365717; PMID:10910347
A;Note: for a complete list of authors see reference number A59328 below
A;Accession: C82521
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1296 <SIM>
A;Cross-references: GB:AE004081; GB:AE003849; NID:g9107985; PIDN:AAF85544.1; GSPDB:GN001
A;Experimental source: strain 9a5c
R;Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, H
as-Neto, E.; Docena, C.; El-Dorry, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A;Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laigr
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E
A;Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;
, F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak
A;Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir
M.; Tsuhako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z
A;Reference number: A59328
A;Contents: annotation
C;Genetics:
A;Gene: XF2759

```
  Query Match             7.1%;  Score 329;  DB 2;  Length 1296;
  Best Local Similarity   21.6%;  Pred. No. 5.1e-09;
  Matches  215;  Conservative  146;  Mismatches  373;  Indels  260;  Gaps   46;

Qy         77 LSLTQTGIAISATKLEKFLQKHSTNKLAKGLDSVENIDRKLGKASNVLSTLSSFLGTALA 136
```

```
Db      178 LDLDGDGIETTGSDGRVILFDHDADGVKTGTGWLKPDDGWLVLDRNGNGTIDS--GRELF 235

Qy      137 GIELDSLIKKGDAAPDAL---------AKASIDLINEIIGNL-----------SQSTQTI 176

Db      236 G--TDTLKRNGQLATDGFDALRDVDSNQDGKIDAADRVFANLRIWRDLNQDGISQANELS 293

Qy      177 EAFSSQLAKLGSTISQAKGFSNIGN-KLQNL--NFSKTNLGLEIITGLLSGISAGF-ALA 232

Db      294 TLDANHIVSIG--VTATAGRVDLGNGNVQTAAGTFTRSN-GTTGATEGTTGTAANLDLLA 350

Qy      233 D--------KNASTGKKVAAGFELSNQVIGNVTKAIS----------SYVLAQRVAAGL- 273

Db      351 DTFYRDFTQQVALTDQAKALPFLRGSGRVRNLDEAISLSKDLGNWVQSYSQQSTRQAQLD 410

Qy      274 ---------STTGAVAALITSSIMLAISPLAFMNAADKFNHANALDE-------FAKQFR 317

Db      411 RLDGLMEKWANTSDMQSLRAQAEALASKGVTVTYVLSGLNPGTAAYQDFLRKLGIVERFM 470

Qy      318 KFGYDGDHLLAEYQRGVGTIEASLTTISTALGAVSAGVSAAAVGSAVGTPIALLVAGV-T 376

Db      471 GFTYGG-------QQG----EARFTPLDATSGTMTVSLSDVQV-----TNIALAYERFKT 514

Qy      377 GLISGILEASKQAMFESVANR--LQGK-ILEWEKQNGGQNYFDKGYDSRYAAYLANNLKF 433

Db      515 DIYESLLLKTRMEPIYNLAEEDFVNGAWVMDW---SGVERALKQGIQRHPRDGILDAIEF 571

Qy      434 LSELNKELEAER------------------------------VIAITQQRW-------- 454

Db      572 VSALGYK-TAERLGWNAIGFLADQLSAAPDMGAFDHESSSWTVIFAAADRHFITGSARSD 630

Qy      455 --------DNNIGELAG----ITKLGERIKSGKAYADAFEDGK---------------K 486

Db      631 VLLGTSGNDMIIGEVTGNDVLIGKGGNDTLQGGSGDDTLDGGTGNDTLNGGAGNDTYRFA 690

Qy      487 VEAG-----SNITLDAKTGIIDISNSNGKKTQALHFTSPLLTAGTESRERLTNGKYSY-- 539

Db      691 IGAGVDSIYEDVTATSDTDIVTFADVRSTALTALERKDYDLVIKYGTSDQLTINNYFHPC 750

Qy      540 -----INKLKFGRVKNWQ--------VTDGEASSKL--DFSKVIQRVAETEGTDEI--GL 582

Db      751 YSGAKIEQFTFSDGVTWDDAAIKARVISNGDASNNYLRGYKDGSNRIYGLDGNDEIYGGA 810

Qy      583 I---VNAKAGNDDIFVGQGKMNIDGGDGHDRVFYSKDGGFGNIT-----------VDGTS 628

Db      811 LDDMLDGGAGNDMLSGGKGADTLDGGSGNDMLY----GGTGNDTYRFAIGAGVDRIEESD 866

Qy      629 ATEAGSYTVNRKVARGDIYHEVVKRQETKVGKRTE--TIQY-------------RDYEL 672

Db      867 AT-AGNTDVVR-------FADVASSALTALERKDSDLVIKYGTSDQLTISNYFYSAEYKV 918

Qy      673 RKVGYGYQST-DNLKSVEEVI--GSQFNDVFKGSKFND---IFHSGEGDDLLDGGAGDDR 726

Db      919 EQFTFSNGVTWDEAAIKARVISNGDANNNYLVG--YNDGSNRIYGLDGNDTIYGGALDDM 976

Qy      727 LFGGKGNDRLSGDEGDDLLDGGSGDDVLNGGAGNDVYIFRKGDGNDTLYDG---TGN-DK 782

Db      977 LYGGAGNDMLSGGKGADTLDGGSGNDMLYGGTGNDTYRFAIGAGVDRIEESDATAGNTDV 1036

Qy      783 LAFADANISDIMIERTKEGIIVKRNDHSGSINIPRWYITSNLQNYQSNKTDHKIEQLIGK 842

Db     1037 VRFADVAFSALTALERKDSDLVIKYGTSDQLTISNYFYSA----------EYKVEQFTFS 1086

Qy      843 DGSYITSDQI-------------DKILQDKKDGT 863

Db     1087 NG--VTWDEAAIKARVISNGDASNNYLRGYKDGS 1118
```

RESULT 28
AE1905
outer membrane secretion protein alr0791 [imported] - Nostoc sp. (strain PCC 7120)
C;Species: Nostoc sp. PCC 7120
A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002
C;Accession: AE1905
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S
DNA Res. 8, 205-213, 2001
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana
A;Reference number: AB1807; MUID:21595285; PMID:11759840
A;Accession: AE1905
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-993 <KUR>
A;Cross-references: GB:BA000019; PIDN:BAB72748.1; PID:g17130136; GSPDB:GN00179
A;Experimental source: strain PCC 7120
C;Genetics:
A;Gene: alr0791

```
  Query Match             6.6%;  Score 305.5;  DB 2;  Length 993;
  Best Local Similarity   22.7%;  Pred. No. 5.6e-08;
  Matches  205;  Conservative  113;  Mismatches  340;  Indels  245;  Gaps   41;

Qy      121 SNVLSTLSSFLGTALAGIELDSLIKKGDAAPDALAKASIDLIN---EIIGNLSQSTQTIE 177

Db       27 TNVGTLVSDIIGNAISDPDASAL--KG------IAVTFVDNSNGTWEYTLNNGTSWTTFG 78

Qy      178 AFSSQLAKLGSTISQAKGFSNIGNKLQ---NLNFSKTN----LGLEIITGLLSGISAGFA 230

Db       79 TPSLTAARLLP--------SNANTKIRFHPNANFSGTADINFYAWDQTTG-TSGSTANI- 128

Qy      231 LADKNASTGKKVAAGFELSNQVIGNVTKAISSYVLAQRVAAG------------LSTTG 277

Db      129 LAGKGGTTA--FSTNYEGASITVTPVNDT------APTLNAGSPTLPTINEDAPLISNKG 180

Qy      278 AVAALITSSIMLAISP----LAFMNAADKFNHA--NALDEFAKQFRKFGYDGDHLLAEYQ 331

Db      181 SLLADLVRGLISDSDPDYQGIA-VTGADNNNGSWQYSLDG-GVNWLNFGNASDNSATLLL 238

Qy      332 RGVGTIEASLTTISTALGAVSAGVSAAAVGSAVGTPIALL--VAGVTGLISGILEASKQA 389

Db      239 PSIRLYDGSLGSLPTSQGWLKFGAS---------PPVFPLFPVGGTQSLITGGIQLNSSS 289

Qy      390 MFESVANRLQGKILEWEKQNGGQNYFDKGYDSRYAAYLAN------------NLKFLSEL 437

Db      290 IGSS----------------GYSNY------NSYAPILFNQAFPELDPVKGFTISFDVKI 327

Qy      438 NKELEAERVIAITQQRWDNNIG--ELAGITKLGERIKSGKAY-------------ADAFE 482

Db      328 NGE---------THTSDDNGDGIQDRAGFSVIVVTSDKTKAIELGFWTDEIWAQTASPLF 378

Qy      483 DGKKVEAGSNITLDAKTGIIDISNSNGKKTQALHFTSPLL-------------TAGTESR 529

Db      379 THSTTERAFRNTTTAVTRYHLVVENNTYKLFAPDSSTPILSGNLRDYSAFNHSTAAPSPI 438

Qy      530 ERLTNGKYSYINKLKFGRVKNWQVTDGEASSKLDFSKVIQRVAETEGTDEIGLIVNAKAG 589

Db      439 TSLPFDPYETPNFLFLGD----NTTSAQASSNL-------TQVELQTNTRVRFVPNAD-- 485

Qy      590 NDDIFVGQGKMNIDGGDGHDRVFYSKDGGFGNITVDGTSATEAGSYTVNRKVARGDIYHE 649

Db      486 ----YNGQANLTFRAWDGSNGVAGGTTG--VNAAVNG-NATAFSSNTLTASITVSPINNP 538

Qy      650 VVKRQETKVGKRTETIQYRDYELRKVGYGYQSTDNLKSVEEVIGSQFNDVFKGSKFNDIF 709

Db      539 I---------------------------QGTTGL---DKLYGTANEDIINGNEGNDYL 566

Qy      710 HSGEGDDLLDGGAGDDRLFGGKGNDRLSGDEGDDLLDGGSGDDVLNGGAGNDVYIFRKGD 769

Db      567 FGRAGNDTLDGGEGNDYLFGGTGNDTLDGGEGSDLLYGNEDNDIINGGVGND--NLDGGT 624

Qy      770 GNDTLYDGTGNDKLAFADANISDIMIERTKEGIIVKRNDHSGSININPRWYITSNLQNY-- 827

Db      625 GDDILRGGTGND--IYTVDTVGDVIEENPNEG-TDKVNSYIS------WTLGANLENLTL 675

Qy      828 --------QSNKTDHKI---------------EQLIGKDGSYIT-----SDQID-KILQD 858
```

```
                                  |: |: |                  : ||||||: |        :|::: :  :|
Db        676 LGNTIIDGTGNELDNHIIGNNAVNRLEGGDGNDWLIGKDGNDILIGGNGNDRLNGETGED 735

Qy        859 KKDGTV--------ITSQELKKLADENKSQKLSASDIASSLN----KLVGSMALFGTANS 906
                 :| :                   : : ||        :|: |   :|        |||: |  |  |
Db        736 TLEGGLGNDVYEIDSVGDVIIEAADAGIDTVISSVDWTLGVNLENLTLVGNQATLGIGND 795

Qy        907 VSS 909
               : :
Db        796 LDN 798
```

RESULT 29
G95851
probable hemolysin-adenlyate cyclase protein [imported] - Sinorhizobium meliloti (strain
C;Species: Sinorhizobium meliloti
C;Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 30-Sep-2001
C;Accession: G95851
R;Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Hernan
Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001
A;Title: The complete sequence of the 1,683-kb pSymB megaplasmid from the N2-fixing endo
A;Reference number: A95842; MUID:21396508; PMID:11481431
A;Accession: G95851
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1072 <KUR>
A;Cross-references: GB:AL591985; PIDN:CAC48479.1; PID:g15139951; GSPDB:GN00167
A;Experimental source: strain 1021, megaplasmid pSymB
R;Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,
pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;
L.; Hyman, R.W.; Jones, T.
Science 293, 668-672, 2001
A;Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,
hebault, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.
A;Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
A;Reference number: A96039; MUID:21368234; PMID:11474104
A;Contents: annotation
C;Genetics:
A;Gene: SMb20079
A;Genome: plasmid

```
  Query Match             6.5%;  Score 301.5;  DB 2;  Length 1072;
  Best Local Similarity   20.7%;  Pred. No. 1e-07;
  Matches  207;  Conservative  140;  Mismatches  346;  Indels  307;  Gaps   43;

Qy         18 TKSGLKNLYLAIPKDYDPQKGGTLNDFIKAADELGIARLAEEPNHTETAKKSVDTVNQFL 77
              | |  :|:|                 ||: || :   |   | | |  |         | ::
Db        152 TISNFENVY-----------GGSGNDILTGDDRSNILR-GEAGNDILNGGADDDLLS--- 196

Qy         78 SLTQTGIAISATKLEKFLQKHSTNKLAKGL---------------DSVENIDRKLGKASN 122
                     |    | :: |   :  |: :   |              |:: |::  :|  ::
Db        197 GGAGNDTADGGTGIDTFDLREKTSSVVVQLSGANAATVFVGGVAEDTIRNVENIVGGTAD 256

Qy        123 VLSTLS-SFLGTALAGIELDSLIKKGDAA--------------PDALAKASIDL-----I 162
                |||          |:|     :  :| |  |              |   |   ::  |       :
Db        257 --DTLSGDAAANKLSGARGNDWLKGGGGADTLDGGEDSDTADYSDKAAAIAVALNGGNPV 314

Qy        163 NEIIGNLSQSTQTIEAFSSQLAKLGSTISQAKGFSNIGNKLQNLNFSKTNLGLEIITGLL 222
                 :| :::              :||: : :   : | : ||:         :   || ::: |
Db        315 TVTVGGIAEDL---------IAKIENIVGGSGGDTIIGDAAANA--FRGGLGADVLDGGG 363

Qy        223 SGISAGFALADK--------NASTGKKVAAGFELSNQV--IGNVTKAISSYVLAQRVAAG 272
                 :| |  :||           | :       | |      :   |  | |:|      :       ||
Db        364 GSDTADF--SDKVQSVVLALNGAVDAIAAVGGTAEDTVRNIENITGGSGNDQFTGDAAAN 421

Qy        273 LSTTGAVAALI---------------TSSIMLAISPLAFMNAADKFNHANALDEFAKQFRK 318
                  |  |  ::                  | |:::   ||  |     |     |  |
Db        422 TFRGGLGADVLDGGDGSDTADYGDKTVSVVVT---LAGANPTTAFVDGMAEDSLRNIENI 478

Qy        319 FGYDGDHLLA--EYQR----GVGTIEASLTTISTALGAVSAGVSAA--AVGSAVGTPIAL 370
                  |  |: :|  :|    :|        | ||  |    :    :   :   |  :  |    ||||   :
Db        479 IGGSGNDVLAGDGFQNVLDGGAGTDMADYSASAKGIAVMLNGANDAKVIVGSAAEDTLR- 537

Qy        371 LVAGVTG-----LISGILEASKQAMFESVANRLQG----KILEWEKQNGGQNYFDKGYDS 421
                : |||      :|:|           :: || | |          ||    | :|||:  | |  :
Db        538 NIENVTGSAFADVITG---------DAQANILLGGSGSDIL---KGDGGQDVIDGGAGT 584

Qy        422 RYAAYLANNLKFLSELNKELEAERVIAITQQRWDNNIGELAGITK------------LGE 469
                | :                     |   |   | |:|:       | |   |   :  |: :                 :|
Db        585 DTADF-----------SEKTAAVVLALAGAA--NAIATVGGLAEDTVRNIESIFGGIGA 630

Qy        470 RIKSGKAYADAFEDGKKVEAGSNITLDAKTGIIDISNSNGKKTQALHFTSPLLTAGTESR 529
                : :|     ::         |    ||:: :||      |:  :    :   ||:::   |    |   |
Db        631 DVLTGDGNSNTIRGG----AGAD-SLDGGAGVDTVDYRD--KTKSVAVTLDGATPVT--- 680

Qy        530 ERLTNGKYSYINKLKFG-----RVKNWQ-VTDGEASSKLDFSKVIQRVAETEGTDEIGLI 583
                                    :| |      ::|:: :: | |    |                |  :   :
Db        681 -------------VKVGGVIEDTIRNFENISGGSAGDMLT-------------GDGLANV 714

Qy        584 VNAKAGNDDIFVGQGKMNIDGGDGHDRVFYSKDGGFGNITVDGTSA-------------- 629
                :     | |  :    | ||    :|||:| |     | ':     . ::|::||::
Db        715 LVGNDGADTLRGGLGKDVLDGGNGVDTADYLEKTDAISVTLNGTASAAVLVGGTAEDTIR 774

Qy        630 ------TEAGSYTVNRKVARGDIYHEVVKRQETKVGKRTETIQYRDYELRKVG------- 676
                       : :|: |:      |    :::      :           |     :|   ||:      ||
Db        775 GVENILSGSGADTLVGDTA-SNMFRGALGADFIDGGAGVDTADYRE----KTGSVEVALS 829

Qy        677 --------YGYQSTDNLKSVEEVIGSQFND---------VFKGSKFNDIFHSGEGDDLLD 719
                         |       | ::::| |  |  :  ||                 |:       |:      | | |:||
Db        830 GASDSFVFVGGVVEDTIRNIENVFGGKGNDTLTGDGLANTLNGNDGKDLLTGGGADILD 889

Qy        720 GGAGDD--------------------------------------RLFGGKGNDRLSGDEGD 742
                |||  |                                     ::|| ||| ||||
Db        890 GGAASDTASYRDKTASVSVTLDGATYTTVTVGGAAEDTIRNVENIWGGTGNDSLSGDANA 949

Qy        743 DLLDGGSGDDVLNGGAGNDVYIFRKGDGN---DTLYDGTGNDKL---------------- 783
                :|| || |  |:| |||| |::  |       |:    ||: | |   |:|
Db        950 NLLSGGGGSDILFGGAGADIFQFDFALGSTNVDTVDLDFTAGDRLFLSKSIFTSLSGGTLA 1009

Qy        784 -----AFADA----NISDIMIERTKEGIIVKRNDHSGSIN 814
                     | |||      |:: :|  |  | :     |  ||::
Db       1010 ATQFYAAADATAAQNVNQKIIYDTTSGALY--YDADGSLS 1047
```

RESULT 30
AG2137
hypothetical protein all2654 [imported] - Nostoc sp. (strain PCC 7120)
C;Species: Nostoc sp. PCC 7120
A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002
C;Accession: AG2137
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi,
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S
DNA Res. 8, 205-213, 2001
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anab
A;Reference number: AB1807; MUID:21595285; PMID:11759840
A;Accession: AG2137
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1417 <KUR>
A;Cross-references: GB:BA000019; PIDN:BAB74353.1; PID:g17131747; GSPDB:GN00179
A;Experimental source: strain PCC 7120
C;Genetics:
A;Gene: all2654

```
  Query Match             6.2%;  Score 288;  DB 2;  Length 1417;
  Best Local Similarity   21.6%;  Pred. No. 7.5e-07;
  Matches  219;  Conservative  131;  Mismatches  355;  Indels  310;  Gaps   50;

Qy          6 VIKSNIQAGLNSTKSGLKNLYLAIP----KDYDPQKGGTLNDFIKA-ADELGIARLAEEP 60
                |   |  |: :    |           |: :         :|||  |  :  | :          ||
Db        525 VFDSGFQSYIEFTAPSDGVYYVGVTLYNNSEYDPNKPASGNGYSSTDPDEYGTGEY---- 580

Qy         61 NHTETAKKSVDTVNQFLSLTQTGIAISATKLEK--------FLQ-----------KHSTN 101
                             |:|  |:    ||    ||::                |||
Db        581 -----------TLNISLN-NPTGFVAGATEIPPGNGAGPTIFLQTIAGTYNNDFENRDFQ 628
```

```
Qy      102 KLAKGLDSVENIDRKLGKASNVL------------------STLSSFLGTALAGIELDS 142
Db      629 ILAPAL--VETSGEDAGSALNIVLVAGGEVPEEGIDVFITSDIVLTDYFG----GLEEDY 682
Qy      143 LIKKG---DAAPDALAKASIDLI----NEIIG---NLSQSTQTI---------EAFSSQ 182
Db      683 SVPYGGNLNGKPFSRGGEFLDAVYNEAGEAIGFRFRLTESFATITLNPTNREEAETDGAE 742
Qy      183 LA------KLGSTISQAKGFSNIG--NKLQ----------NLNFSKTNLGLEIITGLLS 223
Db      743 TATFTIVESLGYQVNPAASSSTVTFYDTLEQVPAPSVTPEVSLAFSTT----ELIESEET 798
Qy      224 GISAGFALADKNASTGKKVAAGFELSNQVIGNVTKAISSYVLAQRVAAG--LSTTGAVAA 281
Db      799 SFTITISLSEPPPAEGVQV--------YISGNAQDALNEFSIFQAQFTGGVPIADGAVSG 850
Qy      282 LI------TSSIMLAISPLAFMNAADKFNHANALDEFAKQFRKFGYDGDHLLAEYQRGVG 335
Db      851 FYFKLLAQTATITLPV-----------FNST----DIVEGIEEFNF-------EVKPGVG 888
Qy      336 --------------------TIEASLTTISTALGAVSAGVSAAAVGSAVGTPIALLVAGV 375
Db      889 YTVNPEQSSAVITIKDTPESEIQVSLSTEPQVL-IETEGTVAKLNFSLSATP---PVAGV 944
Qy      376 TGLISGILEASKQAMFESVANRLQGKILEWEKQNGGQNYFDKGYDSRYAAYLANNLKFLS 435
Db      945 IVTVS-VTDIAENFVTDSL-------VVE-----GGELLGFSAADNTFTLKITEKTASLS 991
Qy      436 ELNKELEAERVIAITQQRWDNNIGELAGITKLGERIKSGKAYADAFEDGKKVEAGSN--- 492
Db      992 -----------VAIA----NDNIAE--GVETATVTLLAG--------DGYQINPLANSGT 1026
Qy      493 ITLDAKTGIIDISNSNGKKTQALHFTSPLLTAGT--------------ESRERL----TN 534
Db     1027 LTLVDSPELAPSSLEESNDTLATAIATGLTTTNTTVTFDSEIAEYTVGEGEERIQVDGTE 1086
Qy      535 GKYSYINKLKFGRVKNWQVTDGEASSKLDFSKVIQRVAETEGTDEIGLIVNAKAGNDDIF 594
Db     1087 DVDLYKVNLKVGEKLSINVDAAEIDSKLLYAQL--RVFDAEGNE------LAKTDFDDFQ 1138
Qy      595 VGQGKMNIDGGDGHDRVFYSKDGGFGNITVDGTSATEAGSYTVNRKVARGDIYHEVVKRQ 654
Db     1139 AAP----------DEVFSAFNDPYLEFTAETT-----GTYYVGISQIGNDYYDPNV--- 1179
Qy      655 ETKVGKRTETIQYRDYELRKVGY---------------GYQSTDNLKSVEE---VIGSQF 696
Db     1180 ---VGSGSGWL-FADFGIENGEYTVSFNLTPEQPTNPVGTSGDDTLIGTDEEESLFGNGG 1235
Qy      697 NDVFKGSKFNDIFHSGEGDDLLDGGAGDDRLFGGKGNDRLSGDEGDDLLDGGSGDDVLNG 756
Db     1236 NDILYARGGDDKLFGGAGDDLLDGGEGNDALFGGAGTDTLLGGAGNDYLTGGTGDNLLDG 1295
Qy      757 GAGNDVYIFRKGD-------GNDTLYDGTGNDKLAFADANISDIMIERTKEGIIVKRNDH 809
Db     1296 GDGNDLLYGNGGQDTLLGGAGDDIIYSGSGDDLINGGLGN-DIIFLNGGQDTIVVAQGAG 1354
Qy      810 SGSINIPRWYITSNLQNYQSNKTDHKIEQLIGKDGSYITSDQ-------IDKILQ 857
Db     1355 IDTIN-----------NFQVS-----LGQKVGLSGG-ITFEQLTFSQSGLDTLIQ 1392
```

RESULT 31
AH2493
hypothetical protein all7128 [imported] - Nostoc sp. (strain PCC 7120) plasmid pCC7120al
C;Species: Nostoc sp. PCC 7120
A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002
C;Accession: AH2493
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S
DNA Res. 8, 205-213, 2001
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana
A;Reference number: AB1807; MUID:21595285; PMID:11759840
A;Accession: AH2493
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-3083 <KUR>
A;Cross-references: GB:BA000020; PIDN:BAB78212.1; PID:g17135666; GSPDB:GN00180
A;Experimental source: strain PCC 7120
C;Genetics:
A;Gene: all7128
A;Genome: plasmid

```
  Query Match             6.2%;  Score 287.5;  DB 2;  Length 3083;
  Best Local Similarity   20.4%;  Pred. No. 2.4e-06;
  Matches  211;  Conservative  141;  Mismatches  348;  Indels  333;  Gaps   45;

Qy        4 INVIKSNIQAGLNSTKSGLKNLYLAIPKDYDPQKGGT----------LNDFIKAADELG 52
Db     2232 VSVVNANGSAG------------SFSYTVSDGKGGTASQTITLEITAVNDAPTLANAIA 2278
Qy       53 IARLAEEPNHTETAKKSVDT-VNQFLSLTQTGIAISATKLEKFLQKHSTNKLAKGL---D 108
Db     2279 NQTATEDTAFTFTIPANTFTDVDAGDALTYSATLADGANLPNWLSFNPSTRTFIGTPTNN 2338
Qy      109 SVENIDRKLGKASNVLSTLSSFLGTALAGIE------LDSLIKKGDAAPDALAKASI--- 159
Db     2339 SVGTVNIRVTATDNAGASVSDVFTLTVANSDTNDAPTLENAIANQTATEDSAFTFTIPAN 2398
Qy      160 --------------------DLINEIIGNLSQSTQTIEAFSSQLAKLGSTISQAKGFSN 198
Db     2399 TFADVDAGDTLTYSATLADGADLLNWL--NFNPSTRTFSGTPTN-DEVGTINIKVTATDN 2455
Qy      199 IGNKLQNLNFSKTNLGLEIITGLLSGISAGFALAD-----------KNASTGKKVAAGF 246
Db     2456 AGASLSDI-FTLTVINTNDAPTVANAIANQTATEDTAFNFQIPADAFNDVDTGDTLTYTA 2514
Qy      247 ELSNQVIGNVTKAISSYVLAQRVAAGLSTTGAVAALITSSIML--------AISPLAFMN 298
Db     2515 TLEN---GDELPSWLTFDAATRTFSGTPTNSEVDTLSIKVIATDKSQASASNVFTLTVLN 2571
Qy      299 AADKFNHANAL-DEFAKQFRKFGY----------DGDHLLAEYQRGVGTIE--ASL---- 341
Db     2572 TNDAPTLENAIADQTATEDSTFSFIIPVNTFADVDADDILA-YS---ATLEEGAALPSWL 2627
Qy      342 -----------TTISTALGAVSAGVSAAAVGSAVGTPIALLVAGVTGLISGILEASKQAM 390
Db     2628 TFNPTNRTFAGTPINSEVGTLNIKVIATDKSSANVSDVFTL------------------ 2668
Qy      391 FESVANRLQGKILEWEKQNGGQNYFDKGYDSRYAAYLANNLKFLSELNKELEAERVIAIT 450
Db     2669 --TVAN----------------------TNDAPILANAI-----ADQAVAANNTFTFT 2697
Qy      451 QQRWDNNIGELAGITKLGERIKSGK--AYADAFEDGKKVEAGSNITLDAKT--------- 499
Db     2698 IP--ENTFSE----------VDTGDILSYSTTLENGDPLPSWLNFNTDTRTFSGNPTTNN 2745
Qy      500 -GIIDI----SNSNGKKTQ---ALHFTSPLLTAGTESRERLTNGKYSYINKLKFGRVKNW 551
Db     2746 AGILNIKVTASDNQGTTVTDIFALTVTASNINPGNDTNNSLS------------------ 2787
Qy      552 QVTDGEASSKLDFSKVIQRVAETEGTDEIGLIVNAKAGNDDIFVGQGKMN-IDGGDGHDR 610
Db     2788 ------GTSSAD-------VLNGFGGDD---YIEGLAGNDTIDGGIGRFDRLFGGDGDDA 2831
Qy      611 V-----FYSKDGGFGNITVDGTSATEAGSYTVNRKVARGDIYHEVVKRQETKVGKRTETI 665
Db     2832 ITDPDGILGAHGGLGNDTINVTFAANWDNDSNPNNSPRSD-------------GKITG-- 2876
Qy      666 QYRDYELRKVGYGYQSTD----------NLKSVEEVIGSQ-FNDVFK--GSKFN------ 706
Db     2877 ----------GYGDDNITVTMNNSKFFINMKGDEPVNNAQGGNDVITLLGSYQNAIVDLG 2926
Qy      707 ---DIFHSGEGDDLLDGGAGDDRLFGGKGNDRLSGDEGDDLLDGGSGDDVLNGGAGNDVY 763
```

```
Db       2927 GGDDTFIGGNGSDNVSGGAGNDTIFGFGGNDNLTGNDGDDILVGGSGNDRLTGGSGKDIF 2986

Qy        764 IFRK-GDGNDTLYD-GTGNDKLAFADANISDIMIERTKEGIIVKRNDHSGSINIPRWYIT 821
              |    || ||: |     :||:     |             |::       :|:::  :: :
Db       2987 SFSSLADGIDTITDFSVADDKIRVNAAGFG--------SGLV------AGNLDASQFVLG 3032

Qy        822 SNLQNYQSNKTDHKIEQLIGKDGSYITSDQIDKILQDKKDGTVITSQELKKLADENKSQK 881
              |: |                 |||       |: : ::  | :          | | :
Db       3033 SSAQ-----------------DGS-------DRFIYNQATGAL--------LFDVDGIGA 3060

Qy        882 LSASDIASSLNKL 894
               :|  ||:  ||:
Db       3061 NTAVQIATLSNKI 3073
```

RESULT 32
T03518
hypothetical protein - Rhodobacter capsulatus
C;Species: Rhodobacter capsulatus
C;Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 08-Oct-1999
C;Accession: T03518
R;Vlcek, C.; Paces, V.; Maltsev, N.; Paces, J.; Haselkorn, R.; Fonstein, M.
Proc. Natl. Acad. Sci. U.S.A. 94, 9384-9388, 1997
A;Title: Sequence of a 189-kb segment of the chromosome of Rhodobacter capsulatus SB1003
A;Reference number: Z14955; MUID:97404404; PMID:9256491
A;Accession: T03518
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-643 <VLC>
A;Cross-references: EMBL:AF010496; NID:g3128256; PIDN:AAC16171.1; PID:g3128319
C;Genetics:
A;Map position: 1

```
  Query Match             6.0%;  Score 280;  DB 2;  Length 643;
  Best Local Similarity   26.1%;  Pred. No. 6.2e-07;
  Matches  110;  Conservative   44;  Mismatches  143;  Indels  124;  Gaps   17;

Qy        409 NGGQNY-FDKGYDSRYAAYLA----NNLKFLSELNKELEAERVIAITQQRWDNNIGELAG 463
              ||    :  | ||:|   |             |:  |        :|:  :   :|
Db         17 NGTSGFSLDPGYESASQAVTITVTDNDAYFSGSATSQLDGNQTAVVTS------------ 64

Qy        464 ITKLGERIKSGKAYADAFEDGKKVEAGSNITLDAKTG----IIDISNSNGKKTQALHFTS 519
                      ||  |              | |:  :   :|    | |:   |   |    : ::
Db         65 --------ASGAVLASG-----AVRLGTATSFTTASGTTARIYDVYVGN---TLVGYVST 108

Qy        520 PLLTAG-----TESRERLTNGKYSYINKLKFGRVKNWQVTDGEASSKLDFSKVIQRVAET 574
               || |      | |    : |                  |   | ||:    | |    :
Db        109 SALTPGVVMTVTGSASATSTG---------------QTYSGLASTSY-ASTVSSALVGG 151

Qy        575 EGTDEIGLIVNAKAGNDDIFVGQGKMNIDGGDGHDRVFYSKD-----GGFGNITVDG--- 626
               | | |       | ||||  :      |    |||| |:||:  :          || || |:||
Db        152 AGNDSI----RAGAGNDTVTGNAGNDTIDGGAGNDRLSGNDGNDLILGGAGNDTLDGGAG 207

Qy        627 --TSATEAGSYTVNRKVARGDIYHEVVKRQETKVGKRTETIQYRDYELR----------- 673
                |  | ||: ::                               : | ||
Db        208 DDTLAGGAGADSIT---------------------GGSGMDYADYSASSGAVTIDLSRW 245

Qy        674 KVGYGYQSTDNLKSVEEVIGSQFND--------VFKGSK-FNDIFHSGEGDDLLDGGAGD 724
                     |  :  | | ||: ||||  :  |          |: ||  :||: | |||  :|| |:
Db        246 SAAGGDAAGDTLTSVDGVIGSAYGDSLIGFDDAVYSGSDVYTNIFYGGGGDDTIDGRGGN 305

Qy        725 DRLF---------GGKGNDRLSGDEGDDLLDGGSGDDVLNGGAGNDVYIFRKGDGNDTLY 775
              | |:            || |:| |  || |:| | ||:| |  |:|||| | ::       | |::
Db        306 DILYGGDDNDLILGGSGDDSLFGDAGNDTLAGGAGADTLSGGAGADTFVLDAQSGADSIT 365

Qy        776 D 776
              |
Db        366 D 366
```

RESULT 33
AC1852
hypothetical protein all0364 [imported] - Nostoc sp. (strain PCC 7120)
C;Species: Nostoc sp. PCC 7120
A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002
C;Accession: AC1852
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S
DNA Res. 8, 205-213, 2001
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anal
A;Reference number: AB1807; MUID:21595285; PMID:11759840
A;Accession: AC1852
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-900 <KUR>
A;Cross-references: GB:BA000019; PIDN:BAB72322.1; PID:g17129709; GSPDB:GN00179
A;Experimental source: strain PCC 7120
C;Genetics:
A;Gene: all0364

```
  Query Match             6.0%;  Score 280;  DB 2;  Length 900;
  Best Local Similarity   19.6%;  Pred. No. 1e-06;
  Matches  188;  Conservative  130;  Mismatches  286;  Indels  354;  Gaps   42;

Qy        118 GKASNVLSTLSSFLGTALAGIEL----------------DSLIKKGDAAPDALAK-ASID 160
              | |::|    : :   |   : |: :                ||:|:            || :  ::|
Db         88 GHATSVAGVIGAAAGNGIGGVGVAYGSTLASFRFSYNNSDSMIR-------ALQRLRNVD 140

Qy        161 LINEIIGNLSQSTQTIEAFSSQLAKLGSTISQAKGFSNIGNKLQN-LNFSKTNLGLEIIT 219
              : |      |: |            | :           |:: :    ::: :  | :  || |:
Db        141 VANNSWGSTS-------IFGADFLN--------PGYAPVVQAIRDAVQFGRNGLGTAIVW 185

Qy        220 GLLSGISAGFALADKNASTGKKVAAGFELSNQVIGNVTKAISSYVLA-QRVAAGLSTTGA 278
                                  | |     |   :     |    ||    |     |:  || |
Db        186 -----------------SAGNSREEGLNTNYSNFSNSRHVISVAALEYDGTASFYSTPG- 227

Qy        279 VAALITSSIMLAISPLAFMNAADKFNHANALDEFAKQFRKFGYDGDHLLAEYQRGVGTIE 338
               |::: |:     : | : :      : :  ::| :: :|
Db        228 -ASILVSAFGSGV-PGSIVTTDRRGSEGSSLGDYNYEFN--------------------- 264

Qy        339 ASLTTISTALGAVSAGVSAAAVGSAVGTPIALLVAGVTGLISGILEASKQAMFESVANRL 398
                                | ||||                |:|::: :|||::   :   :    |
Db        265 ---------------GTSAAA--------------PEVSGVVALMLEANRNLGYRDIQEIL 296

Qy        399 QGKILEWEKQNGGQNYF------------------DKGY---DSRYAAYLANNLKFLSEL 437
                  : :  | | ||                     | |:      |:   |   ||    :  |
Db        297 AYSARQNDFYNVGGNYIWQINGANNFNGGGLHVSHDYGFGLVDALAAVRLAETWQKQSRF 356

Qy        438 NKELEAERVIAITQQRWDNNIGELAGITKLGERIKSGKAYADAFEDGKK---VEAGSNIT 494
              | |             |        : |  | |:| : :  ::| ::       |  :      ||  |:|
Db        357 NNE-----------QSLSYSSGNL-GLT-VPDNDEAGISHTFTVAAGLEIDWVEVELNLT 403

Qy        495 LDAKTGIIDISNSNGKKTQALHFTSPLLTAGTESRERLTNGKYSYI-----NK------- 542
                 :  |:               :: |||               :|   |  :      ||
Db        404 HPYRGDIV-------------VYLTSP-------------SGVQSVLVHQPGNKEDEGDNI 438

Qy        543 -LKFGRVKNWQVTDGEASSKLDFSKVIQRVAETEGTDEIGLIVNAKAGNDDIFVGQGKMN 601
                |       ::|     || |:  :::  ||  :        |  :||:                |:|
Db        439 VFKLSSTQHW----GETSAG-NWTLTIQDL----GPSDIGIF------------NSWKLN 477

Qy        602 I--DGGDGHDRVFYSKDGGF-GNITVDGTSATEAGSYTVNRKVARGDIYHEVVKRQETKV 658
              :   |    :|  ||: : || |: |:  :| |:       |:|         | |  :     : :
Db        478 LYGDADTINDTYFYTNEYGFYGSTTLTDSSGTD----TINAAAITADSYLNLNPGSTSIL 533

Qy        659 GKRTETIQYRDYELRKVGYGYQSTDNLKSVEEVIGSQFNDVFKGSKFNDIFHSGEGDDLL 718
                ||                  ||      ::|    |  ||   :    ::  | :  |:| |
Db        534 NGTVLTI----------------STGT--TIENAFGGDGNDTIIGNSAANVLHGGRGNDTL 576

Qy        719 DGGAGDDRLFGGKGND----RLSGD-------EGDD------------------------ 743
              ||| |:| |  ||:|||         :||          || |
Db        577 DGGVGNDTLRGGRGNDTYIVNSTGDIVTENANEGIDTVQSSVTYTLGANVENLTLTGTGA 636

Qy        744 ----------LLDGGSGDDVLNGGAGNDVYIFRKGDGNDTLYDGTGNDKLAFADAN---- 789
                          : | ||::  ||| ||||  |    |:||| |    |||||  :       |
Db        637 INGTGNSLNNTITGNSGNNTLNGDAGNDFLI--AGNGNDILNGGTGNDTMLGGGGNDTYI 694
```

```
Qy        790 ---ISDIMIERTKEGIIVKRNDHS------------------GSINIPRWYITSNLQNY 827
                 | | ::|   :|  : ::  |                   |: |     || |  |
Db        695 VDSIGDYVLENANQGTDLVQSSISYTLGNSLENLTLTGTSAINGTGNRLNNVITGNSGNN 754

Qy        828 QSNKTDHK--------IEQLIGKDGSYI----------------------TSDQIDKIL 856
                 |  |             :: |:|  :|: |                    |: ||:|
Db        755 TLNGGDGNDTLNGSAGVDTLLGGNGNDILVGGTGNDTLTGGVGRDRFTFNSRSEGIDRIT 814

Qy        857 Q-DKKDGTVITSQELKKLADENKSQKLSASDIASSL---NKLVGSMALFGTANSVSSN 910
                 :  | |::                 :||:     |       :   |  | |:| : :|:
Db        815 DFNVVDDTIV----------------VSAAGFGGGLVVGAAIASSQFLLGSAATTASH 856
```

RESULT 34
G87572
calcium-binding protein [imported] - Caulobacter crescentus
C;Species: Caulobacter crescentus
C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 20-Apr-2001
C;Accession: G87572
R;Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A;Title: Complete Genome Sequence of Caulobacter crescentus.
A;Reference number: A87249; MUID:21173698; PMID:11259647
A;Accession: G87572
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-650 <STO>
A;Cross-references: GB:AE005673; NID:g13424183; PIDN:AAK24579.1; GSPDB:GN00148
C;Genetics:
A;Gene: CC2610

```
  Query Match             5.9%;  Score 276;  DB 2;  Length 650;
  Best Local Similarity   26.0%;  Pred. No. 1e-06;
  Matches  107;  Conservative   51;  Mismatches  118;  Indels  136;  Gaps   17;

Qy        588 AGNDDIFVGQGKMNIDGGDGHDRVFYSKDGGFGNITVDGTSATEAGSYTVNRKVARGDIY 647
              ||:| |    :|   : |||| | :       || || |:|| : :: :|
Db         34 AGDDAIAGHEGNDVLQGGDGDDYML----GGPGNDTLDGGAGSDWAAY----------- 77

Qy        648 HEVVKRQETKVGKRTETIQYRDYEL---RKVGYGYQSTDNLKSVEEVIGSQFNDVFKGSK 704
                 ::     | :          |  |     :  | |     || | |:| : || |||    |:
Db         78 ------EDATAGVKV------DLNLSGAQNTGGG--GTDRLTSIENLYGSAFNDTLIGNA 123

Qy        705 FNDIFHSGEGDDLLDGGAGDDRLFGGKGNDRLSGDEGDDLLDGGSGDDVLNGGAGND--- 761
               :::   | | |  :  || |||  |:|   ||| |  | :||| : ||:||||: ||||  |
Db        124 GDNMITGGAGADSITGGKGDDTLWGSAGNDTLDGGDGDDWMVGGAGDDVIKGGAGGDWSS 183

Qy        762 ------------------------------------------------VYIFRKGD-GN 771
                                                              ||:  || |:
Db        184 YEDATAGVTVDLTKTTAQDTIGAGKDTLSGVENLWGSKFADVLTGDAGSNYIW--GDAGD 241

Qy        772 DTLYDGTGNDKLA-FADANISD-------IMIERTKEGIIVKRNDHSGSINIPRWYITSN 823
              | ||| | |:|  |  |  |  |: |          |     :  |:         ::: |   ||
Db        242 DKLYGGAGDDFFAGGAGVNVIDGGEGFDTIDYGMSAVGV---------EVDLSRNIATSR 292

Qy        824 LQNYQSNKTDHKIEQLIGK----------DGSYITSDQIDKILQDKKDGTVITSQELKKL 873
               :    : |     || : |               :|:  |  : :|: :  | |:  |
Db        293 FGDTSISDTLSSIELVTGSIYADFIIGNMAENYLYGDAGNDVLRAEGGGDVLEGGE---- 348

Qy        874 ADENKSQKLSASDIASSLNKLVG-----SMALFGTANSVSS----NALQPIT 916
                |:      |  |||        |:|       ::: | :|         :|||  |
Db        349 GDDFLIGSL--SDI------LIGGGGDDQLSVSGGSNYVDGGDGVDALQTFT 392
```

RESULT 35
H95964
probable outer membrane secretion protein SMb21543 [imported] - Sinorhizobium meliloti (
C;Species: Sinorhizobium meliloti
C;Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 30-Sep-2001
C;Accession: H95964
R;Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Hernan
Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001
A;Title: The complete sequence of the 1,683-kb pSymB megaplasmid from the N2-fixing endo
A;Reference number: A95842; MUID:21396508; PMID:11481431
A;Accession: H95964
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1112 <KUR>
A;Cross-references: GB:AL591985; PIDN:CAC49384.1; PID:g15140870; GSPDB:GN00167
A;Experimental source: strain 1021, megaplasmid pSymB
R;Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,
pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;
L.; Hyman, R.W.; Jones, T.
Science 293, 668-672, 2001
A;Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,
hebault, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.
A;Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
A;Reference number: A96039; MUID:21368234; PMID:11474104
A;Contents: annotation
C;Genetics:
A;Gene: SMb21543
A;Genome: plasmid

```
  Query Match             5.8%;  Score 269.5;  DB 2;  Length 1112;
  Best Local Similarity   24.5%;  Pred. No. 4.7e-06;
  Matches  127;  Conservative   56;  Mismatches  143;  Indels  193;  Gaps   24;

Qy        407 KQNGGQNYF---DKGYDSRYAAYLANNLKFLSELNKELEA--ERVIAI-------TQQRW 454
               :|    :||   |  |  ||::|           : |    ::||    |   :            | |
Db        183 QQYSVMSYFDGIDVGETSRWSA--------STPLMADIEAVIRRFFSTVDENGVRTYQHI 234

Qy        455 DNNIGELA---GITKLGERIKSGKAYADAFEDGKKVEAGSNITLDAKTGIIDISNSNGKK 511
               | | |:      | |: | :: |             | : :  | .|           ||| | |
Db        235 DLNTGDNVYGFGSTQYGYQLTS---------SGMQHDIGFAIHDTGGVDTIDFSGS---- 281

Qy        512 TQALHFTSPLLTAGTESRERLTNGKYSYINKLKFGRVKNWQVTDGEASSKLDFSKVIQRV 571
                          ||||    |   |::|  :|    |     |    :   |    :    |:
Db        282 -----------TAGTIL--DLRAGQFSSVN----GHSNNVSIFAGHNADAADY------- 317

Qy        572 AETEGTDEIGLIVNAKAGN--DDIFVG-QGKMNIDGGDGHDRVFYSKDGGFGNITVDGTS 628
                           :     |:   |||  :|   |    :||    |||:    :  :||     ||
Db        318 -----------YIETGIGSRFDDILIGNDGANTLDGRSGGDRM--AGNGGDDTYFVDSL- 363

Qy        629 ATEAGSYTVNRKVARGDIYHEVVKRQETKVGKRTETIQYRDYELRKVG------YGYQST 682
                                   | :  |:|       |      |:  ::||:         :||
Db        364 -------------------EDIVREEANGGNDTVILLSRNLKIRKIANVENIIYADEST 403

Qy        683 --------------DN-LKSVEEVIGSQFNDVFKGSKFNDIFHSGEGDDLLDGGA----- 722
                            || :  |:   :     ||     |     :|       ||||: ||
Db        404 TQPGGGGASPPSAGDNTMTSI--IFGGNGNDTLDGGAGDDTIFGKGGDDLIIGGRDSLAS 461

Qy        723 -------------------GDDRLFGGKGNDRLSGDEGDDLLDGGSGDDVLNGGAGNDV 762
                                 |:| |:|| ||| :  :|:|:|||||:|||||:|   | |
Db        462 RDINNTIDVEDLEDQTESDDGNDTLYGGGGNDTILGGQGNDILDGGAGDDVLSGQDGVD- 520

Qy        763 YIFRKGDGNDTLYDGTGNDKLAFADANISDIMIERTKEG-----IIVKRNDHSGSINIPR 817
               ||| | | ||                                  :: :||        : :   |  ||
Db        521 -IFRGGAGVDT--------------------VDFSKESPFQLLVNLATNVASGGT---- 554

Qy        818 WYITSNLQNYQSNKTDHKIEQLIGKDGSYITSDQIDKIL 856
                         |  |  ||  ||| |        |:||: :
Db        555 ----------ASGDTFHSIENLIGSD------DRIDRFI 577
```

RESULT 36
G95405
hypothetical protein SMa2111 [imported] - Sinorhizobium meliloti (strain 1021) magaplasm
C;Species: Sinorhizobium meliloti
C;Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 30-Sep-2001
C;Accession: G95405
R;Barnett, M.J.; Fisher, R.F.; Jones, T.; Komp, C.; Abola, A.P.; Barloy-Hubler, F.; Bows
.; Kalman, S.; Keating, D.H.; Palm, C.; Peck, M.C.; Surzycki, R.; Wells, D.H.; Yeh, K.C.
Proc. Natl. Acad. Sci. U.S.A. 98, 9883-9888, 2001

```
A;Title: Nucleotide sequence and predicted functions of the entire Sinorhizobium melilot
A;Reference number: A95262; MUID:21396509; PMID:11481432
A;Accession: G95405
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-539 <KUR>
A;Cross-references: GB:AE006469; PIDN:AAK65809.1; PID:g14524311; GSPDB:GN00165
A;Experimental source: strain 1021, megaplasmid pSymA
R;Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,
pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;
 L.; Hyman, R.W.; Jones, T.
Science 293, 668-672, 2001
A;Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,
hebault, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.
A;Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
A;Reference number: A96039; MUID:21368234; PMID:11474104
A;Contents: annotation
C;Genetics:
A;Gene: SMa2111
A;Genome: plasmid

  Query Match             5.8%;  Score 268;  DB 2;  Length 539;
  Best Local Similarity   26.4%;  Pred. No. 2e-06;
  Matches  103;  Conservative   59;  Mismatches  160;  Indels   68;  Gaps   13;

Qy        519 SPLLT--AGTESRERLTNGKYSYINKLKFGRVKNWQVTDGEASSKLDFSKVIQRVAETEG 576

Db         21 SPIMTRIVGTSADDTLDG-----------------MAEGDRIWSLDGNDVVDGGA---G 59

Qy        577 TDEIGLIVNAKAGNDDIFVGQGKMNIDGGDGHDRVFY-----SKDGGFGNITVDGTSATE 631

Db         60 DD----FVDGGAGDDALTSSSGFDEFTGGEGNDRLSFIGVGGAARGGTGVDTLVGDYAAI 115

Qy        632 AGSYT---VNRKVARGD-------IYHEVVKRQETKVGKRTETIQYRDYELRKV--GYGY 679

Db        116 SDAFLFDGMHGHAAFGDLSVKGNHLYFLDIERLSLTTGIGDDRIIATGFSFVHVHTGAGD 175

Qy        680 QSTDNLKSVEEVIGSQFNDVFKGSKFNDIFHSGEGDDLLDGGAGDDRLFGGKGNDRLSGD 739

Db        176 DRVETGIGDDQIYAGDGRDLLFGGGDDFIHTGQGDDYVNGGNDDDRLEGEDGNDSLVGG 235

Qy        740 EGDDLLDGGSGDDVLNGGAGNDVYIFRKGDGNDTLYDGTGNDKL--AFADANI------S 791

Db        236 RGNDRLDGGSGDDDVNGGDGND--SLTGGLGSDTVTGGAGDDYLSNSFAAGDILLGGDGN 293

Qy        792 DIMIERTKEGIIVKRNDHSGSINIPRWYITSNLQNYQSNKTDHKIEQLIGKDGSYITSDQ 851

Db        294 DTLSAGGEDTAYGGWSDLYGGAGDDRLHI----------YTDGSLGALDGGEG----FDR 339

Qy        852 IDKILQDKKDGTVITSQELKKLADENKSQK 881

Db        340 ASIALDDVSAGFVLDASRFGSIEEFNITVK 369


RESULT 37
E95995
hypothetical calcium binding protein [imported] - Sinorhizobium meliloti (strain 1021) m
C;Species: Sinorhizobium meliloti
C;Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 30-Sep-2001
C;Accession: E95995
R;Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Hernan
Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001
A;Title: The complete sequence of the 1,683-kb pSymB megaplasmid from the N2-fixing endo
A;Reference number: A95842; MUID:21396508; PMID:11481431
A;Accession: E95995
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-387 <KUR>
A;Cross-references: GB:AL591985; PIDN:CAC49629.1; PID:g15141116; GSPDB:GN00167
A;Experimental source: strain 1021, megaplasmid pSymB
R;Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,
pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;
 L.; Hyman, R.W.; Jones, T.
```

```
Science 293, 668-672, 2001
A;Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,
hebault, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.(
A;Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
A;Reference number: A96039; MUID:21368234; PMID:11474104
A;Contents: annotation
C;Genetics:
A;Gene: SMb21402
A;Genome: plasmid

  Query Match             5.8%;  Score 267.5;  DB 2;  Length 387;
  Best Local Similarity   30.8%;  Pred. No. 1.3e-06;
  Matches   89;  Conservative   39;  Mismatches   92;  Indels   69;  Gaps   15;

Qy        545 FGRVKNWQVTDGEASSKL------DFSKVIQRVAETEGTDEI--GL------------IV 584

Db         31 YGGVGNDRIFGGDGNDYIVGGVGDDYAE------GGEGNDRLDGGLGNDAFMGGIGNDIF 84

Qy        585 NAKAGNDDIFVGQGKMNIDGGDGHDRVFYSKDGGFGNITVDGTSATE-----AGSYTVNR 639

Db         85 DGGEGNDVLDGGHGHDQMLGGRGHDKIF----GGAGEDYIDADEGDDIIFAGSGDDGFNN 140

Qy        640 KV--ARGDIYHEVVKRQETKVGKRTETIQYRDYELRKVGYGYQSTDNLKSV---EEVIGS 694

Db        141 RVNPATGKLTQQAVGG-----GAGNDTI-----------YGEGGNDALKGQSGNDRVYGG 184

Qy        695 QFNDVFKGSKFNDIFHSGEGDDLLD---------GGAGDDRLFGGKGNDRLSGDEGDDLL 745

Db        185 IGDDIVDGGNGNNYLDGGDGNDVLDSEGGSDEAHGGSGNDRISVGAGNDRAFGDDGDDIL 244

Qy        746 DGGSGDDVLNGGAGNDVYIFRKGDGNDTLYDGTGNDKLAFADANISDIM 794

Db        245 TSGAGDDYLSGGIHQD-RLF-GGTGNDTLRGDAGRDVLS-GDAG-SDIL 289


RESULT 38
AD2263
hypothetical protein alr3659 [imported] - Nostoc sp. (strain PCC 7120)
C;Species: Nostoc sp. PCC 7120
A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002
C;Accession: AD2263
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi,
 Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S
DNA Res. 8, 205-213, 2001
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anab
A;Reference number: AB1807; MUID:21595285; PMID:11759840
A;Accession: AD2263
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-589 <KUR>
A;Cross-references: GB:BA000019; PIDN:BAB75358.1; PID:g17132792; GSPDB:GN00179
A;Experimental source: strain PCC 7120
C;Genetics:
A;Gene: alr3659

  Query Match             5.7%;  Score 266.5;  DB 2;  Length 589;
  Best Local Similarity   25.7%;  Pred. No. 2.7e-06;
  Matches  114;  Conservative   64;  Mismatches  147;  Indels  119;  Gaps   21;

Qy        411 GQNYFDKGYDS-RYAAYLANNLKFLSELNKELEAERVIAITQQRWDNNIGELAGITKLGE 469

Db         36 GDDYLYGGLDNDKLYGGLGNDKLYGGEGNDNLNGE--------QGNNSLYGEAGDDTLDA 87

Qy        470 RIKSGKAYADAFEDGKKV---EAGSNITLDAKTG--IIDISNSNGKKTQALHFTSPLLTA 524

Db         88 GNSTGNNLLDGGEGNDEIYVYSSSGNNTLKGGVGNDILDVRFSYGKN---------ILDG 138

Qy        525 GTESRE---RLTNGKYSYINKLKFGRVKNWQVTDGEASSKL----DFSKVIQRVAETEGT 577

Db        139 GSGNDQLYARYSQGN----NTLRGGSGDDYLGVDLSAGNNVLNGDDGNDILDVSYQWKGK 194

Qy        578 DEIGLIVNAKAGNDDIFVG---QGKMNIDGGDGHDRVFYSK---------------DGGF 619
```

```
                :    ||:   || |||      ||    ::|| | |   : |                       |||
Db        195 N----IVSGGNGN-DIFYAYGVQGANTLNGGSGDDSFYISSPSADTVVHPLVTQTVDGGI 249

Qy        620 GN--ITVDGTSATEAGSYTVNRKVARGDIYHEVVKRQETKVGKRTETIQYRDYELRKVGY 677
              || : :| ::|      : : |     :| |                    :| | |
Db        250 GNDYLYIDYSTAIAGITSSFNTTTKQGLI--------------TADTNQVR--------- 286

Qy        678 GYQSTDNLKSVEEVIGSQFNDVFKGSKFNDIFHSGE-GDDLLDGGAGDDRL--------- 727
               |:: : |      :||:  :|:  |    :||    |   |:| |:||||:| |
Db        287 -YKNIERL----NIIGTSHDDIIVGGNSDDIIDGGSGGNDTLNGGAGNDILIIRDNSRSH 341

Qy        728 ---FGGKGNDRL--SGDEGDDLLDGGSGDDVLN----------------GGAGNDVYIFR 766
                 :|| ||| |  :|   | :|||||  |:| |:               ||||:|
Db        342 NTVYGGAGNDSLYANGTAGTNLLDGGDGNDYLSAKNNYSYFFNASQTLIGGAGDDTLDVS 401

Qy        767 KGDGNDTLYDGTGNDKLAFADANI 790
                ||: || | ||| | :|| ::
Db        402 GSYGNNFLYGGAGNDFL-YADYSV 424
```

RESULT 39
AD1841
hypothetical protein alr0276 [imported] - Nostoc sp. (strain PCC 7120)
C;Species: Nostoc sp. PCC 7120
A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002
C;Accession: AD1841
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S
DNA Res. 8, 205-213, 2001
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana
A;Reference number: AB1807; MUID:21595285; PMID:11759840
A;Accession: AD1841
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-2348 <KUR>
A;Cross-references: GB:BA000019; PIDN:BAB77800.1; PID:g17135254; GSPDB:GN00179
A;Experimental source: strain PCC 7120
C;Genetics:
A;Gene: alr0276

```
  Query Match             5.7%;  Score 266;  DB 2;  Length 2348;
  Best Local Similarity   20.4%;  Pred. No. 2.1e-05;
  Matches  193;  Conservative  121;  Mismatches  332;  Indels  298;  Gaps   38;

Qy         16 NSTKSGLKNL----YLAIPKDYDPQKGGTLNDFIKAADELGIARLAEEPNHTETAKK--S 69
              ||   |:: :         | |          | ||| |      ::        |:  | : |::
Db       1479 NSAFGGIRTIEGVVVAAFP------GGSGLNGFFVQEED------ADTDNDSTTSEGIFV 1526

Qy         70 VDTVNQFLSLTQTGIAISATKLEKFLQKHSTNKLAKGLDSVENIDRKLGKASNVLSTLSS 129
                | ||      :|     ::   :  : |||      | : ::    :  :::| :|:
Db       1527 FDPTGQF-----SGSVGDKVRVTGSVSEFSTNNGVSSLTQLSSVSSIINLGADILPKMSN 1581

Qy        130 FL--GTALAGIELDSLIKKGDAAPDALAKASIDLINEIIGNLSQSTQTIEAFSSQLAKLG 187
                | :| :|·   ::                 :|    ||        | |  | |  || : |
Db       1582 IQLPVTTVADLERYEGMR----------------VNISAGN-GDLTVT-EHF--QLGRFG 1621

Qy        188 STISQAKGFSNIGNKLQNLNFSKTNLGLEIITGLLSGISAGFA-----------LADKNA 236
                :  | | ||                |:  ||  |          ||:|          : |  :
Db       1622 QVVLSATGTSNQPG---------TDGRLEQYTQFNDPSVAGYAAYLDEIAKRRIILDDGS 1672

Qy        237 ST-----------GKKVAAGFELSNQVIGNVTKAISSYVLAQR-------VAAGLSTTGA 278
              ||           |: ::|       :| : |  | |  : || ||            : |: | |
Db       1673 STQNPATIIFGRGGEPLSA----TNTLRGGDTVASITGVLDQRFEGYRVQTSTGVDFTPA 1728

Qy        279 VAALITSSIMLAISPLAFMNAADKFN--------------HANALDEFAKQFRKF----- 319
                      |:  :      :|  |   : ||                  |  | || :|  |
Db       1729 NPRPATTPDVGGTLKVASFNVLNYFNGDGTGSGFTSPEQRGAENLTEFNRQREKTIAAIL 1788

Qy        320 ---------------GYDGDHLLAEYQRGVGTIEASLTTISTALGAVSAGVSAAAVGSAV 364
                             || :  : :    |: : :    :      |    | |||  :
Db       1789 GLNADVVGLIEIENDGYGANSAIQDLINGLNAVAGAGTYAFINPGLSQLGTDAIAVG-FI 1847

Qy        365 GTPIALLVAGVTGLISGILEASKQAMFESVANRLQGKILEWEKQNGGQNYFDKGYDSRYA 424
               | ::   ||    ::                                           || ||
Db       1848 YKPNSVTPVGVAATVA--------------------------DGFGQGAFDN------- 1873

Qy        425 AYLANNLKFLSELNKELEAERVIAITQQRWDNNIGELAGITKLGERIKSGKAYADAFEDG 484
                   || | |:: ::                           |: ||     |  :      ||
Db       1874 ----NNRKPLAQTFRQ---------------NSTGE--QFTAVINHFKS----------- 1901

Qy        485 KKVEAGSNITLDAKTGIIDISNSNGKKTQALH------FTSPLLTAGTE----------- 527
                 | :|:     ||  |          ||| :|:|          |:|  |   |:
Db       1902 KGSSSGNPGDADAGDG---QGLSNGTRTRASQDLAAWLATNPTGTTDTDYLILGDLNAYA 1958

Qy        528 -----------SRERLTNGKYSYINKLKFGRVKN----------------WQVTDGEAS 559
                              | |   |||:    ::| : :                    | : : :
Db       1959 QEDPIRALENAGYNNLLPNTTYSYVFDGQWGALDHALANASLASQVSTAVKWHI-NADEP 2017

Qy        560 SKLDFSKVIQRVAET---------EGTDE----IGLIVNAK--AGND---------DIF 594
               : ||::     : |  :              :|          :|| :|      | ||             :|
Db       2018 NVLDYNTNFKSVGQQTSLYSPDAFRSSDHDPVIVGLNLNTAPIAVNDIATTNENTAVNIN 2077

Qy        595 VGQGKMNIDGGDGHDRVFYSKDGGFGNITVDGTSATEAG---SYTVNRKVARGDIYHEVV 651
              |     :::|      :     :       |        :   :||   |  |  |  |     :    :
Db       2078 VLTNDSDVNGDSLQLSLVSNPVNGVAVVNDNGTPGNFADDFITYTPNTGYLGNDSFTYSI 2137

Qy        652 KRQETKVGKRTETIQYRDYELRKVGYGYQSTDNLKSVEEVIGSQFNDVFKGSKFNDIFHS 711
               :  |  |    | |:                       :  :   :||:   ||:  |:   ||:
Db       2138 --SDGKGGTATATVSL----------------TINASGGIIGTPDNDILTGTNRNDLIRG 2179

Qy        712 GEGDDLLDGGAGDDRLFGGKGNDRLSGDEGDDLLDGGSGDDVLNGGAGNDVYIFRKGDGN 771
               |:||| || |:| |:| :| | | |   |:| | ||   :| | | || |:|: :   ||
Db       2180 LGGNDLLIGGNGNDTLYGDRGKDILLGGNGNDTLYGGDDNDTLIGGNGDDLLV--GGKGN 2237

Qy        772 DTLYDGTGNDKLAFADAN------ISD-------IMIERTKEGI 802
              | |  | | |:    :|            |:|        |:| | : |:
Db       2238 DLLIGGNGRDRFYLSDTRTGEFDIITDFKVGQDTILIPRAEFGL 2281
```

RESULT 40
E95933
probable calcium-binding exported protein [imported] - Sinorhizobium meliloti (strain 10:
C;Species: Sinorhizobium meliloti
C;Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 30-Sep-2001
C;Accession: E95933
R;Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Hernan
Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001
A;Title: The complete sequence of the 1,683-kb pSymB megaplasmid from the N2-fixing endo
A;Reference number: A95842; MUID:21396508; PMID:11481431
A;Accession: E95933
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-387 <KUR>
A;Cross-references: GB:AL591985; PIDN:CAC49133.1; PID:g15140618; GSPDB:GN00167
A;Experimental source: strain 1021, megaplasmid pSymB
R;Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,
pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;
L.; Hyman, R.W.; Jones, T.
Science 293, 668-672, 2001
A;Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,
hebault, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.
A;Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
A;Reference number: A96039; MUID:21368234; PMID:11474104
A;Contents: annotation
C;Genetics:
A;Gene: SMb21229
A;Genome: plasmid

```
  Query Match             5.7%;  Score 263.5;  DB 2;  Length 387;
  Best Local Similarity   25.4%;  Pred. No. 2.1e-06;
  Matches   99;  Conservative   51;  Mismatches  111;  Indels  129;  Gaps   15;

Qy        546 GRVKNWQVTDGEASSKLDFSKVIQRVAETEGTDEIGLIVNAKAGNDDIFVGQGKMNIDGG 605
               | : |  :  |: |  ||             :| |:|           |:| || | |:  || |
Db         77 GGLGNDILDGGDGSDILDGG---------DGHDQI----LGGRGHDKIFGGAGEEYIDAG 123
```

```
Qy   606 DGHDRVFY-SKDGGFGNITVDGTSATEAGSYTVNRKVARGDIYHEVVKRQETKVGKRTET 664
         || | ::  | | || |       |         | ::   |    |: :
Db   124 DGDDIIYAGSGDDGFNNRVNPATG-------QVTQQAVAGGAGHDKI------------- 163

Qy   665 IQYRDYELRKVGYGYQSTDNLKSV---EEVIGSQFNDVFKGSKFNDIFHSGEGDDLLD-- 719
                     || :  | ||       : | |    ||:  |    |:    |:|:|:||
Db   164 ------------YGEEGNDALKGQSGNDWVYGGIGNDIVDGGDGNNYLDGGDGNDVLDSE 211

Qy   720 -------GGAGDDRLFGGKGNDRLSGDEGDDLLDGGSGDDVLNGGAGNDVYIFRKGDGND 772
                ||  |:|::  | |:||: ||:|||:| | :|:| |:||    |  ||  |||||
Db   212 GGIDEAHGGTGNDKIVVGAGDDRVFGDDGDDMLYGEAGNDSLSGGNNYD-RIF-GGDGND 269

Qy   773 T-----------------LYDGTGNDKLAFADA-------------NISDIMIERTKE 800
         |                 |: |  :|:  |  |             :  ::||:
Db   270 TLRGDAGRDVLIGEAGTDILWGGADSDRFVFKGAPVLSGQDTVMDFQDGVDFLVIEK--- 326

Qy   801 GIIVKRNDHSGSINIPRWYITSNLQNYQSNKTDHKIEQLIGKDGSYITSDQIDKILQDKK 860
          : :|:   ||:        :  | :   |  ::         | :| :    || |
Db   327 -LGIKQYSSSGA--------AGTVYAYDATGGDVLVK-------GYDSSGKAFSILLDDP 370

Qy   861 DGTVITSQELKKLADENKSQKLSASDIASS 890
          :||                 |||:: :||
Db   371 NGT-----------------LSAANFSSS 382
```

Search completed: February 17, 2004, 10:12:18
Job time : 40 secs

```
                    GenCore version 5.1.6
          Copyright (c) 1993 - 2004  Compugen Ltd.

OM protein - protein search, using sw model

Run on:         February 17, 2004, 10:09:49 ; Search time 45 Seconds
                                          (without alignments)
                                          5315.884 Million cell updates/sec

Title:          US-10-069-799-5
Perfect score:  4647
Sequence:       1 MSNINVIKSNIQAGLNSTKS..........SSNALQPITQPTQGILAPSV 927

Scoring table:  BLOSUM62
                Gapop 10.0 , Gapext 0.5

Searched:       830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters:      830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       SPTREMBL_23:*
                1:  sp_archea:*
                2:  sp_bacteria:*
                3:  sp_fungi:*
                4:  sp_human:*
                5:  sp_invertebrate:*
                6:  sp_mammal:*
                7:  sp_mhc:*
                8:  sp_organelle:*
                9:  sp_phage:*
                10:  sp_plant:*
                11:  sp_rodent:*
                12:  sp_virus:*
                13:  sp_vertebrate:*
                14:  sp_unclassified:*
                15:  sp_rvirus:*
                16:  sp_bacteriap:*
                17:  sp_archeap:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                                 SUMMARIES
                %
Result          Query
  No.   Score   Match Length DB  ID                  Description
-------------------------------------------------------------------------
   1     4642   99.9    927  2  Q93GI2               Q93gi2 moraxella b
   2   2358.5   50.8    946  2  Q9EV24               Q9ev24 mannheimia
   3     2356   50.7    953  2  Q9EV23               Q9ev23 mannheimia
   4     2355   50.7    953  2  Q9ETX2               Q9etx2 mannheimia
   5     2355   50.7    953  2  Q9EV25               Q9ev25 mannheimia
   6     2345   50.5    953  2  Q9EV26               Q9ev26 mannheimia
   7   2338.5   50.3    953  2  Q9EV32               Q9ev32 pasteurella
   8   2336.5   50.3    953  2  Q9EV33               Q9ev33 pasteurella
   9   2332.5   50.2    953  2  Q9EV28               Q9ev28 pasteurella
  10     2332   50.2    953  2  Q9ETG5               Q9etg5 pasteurella
  11   2329.5   50.1    953  2  Q9EV31               Q9ev31 pasteurella
  12   2329.5   50.1    953  2  Q9EV27               Q9ev27 pasteurella
  13     2329   50.1    953  2  Q9EV29               Q9ev29 pasteurella
  14     2326   50.1    953  2  Q9EV34               Q9ev34 pasteurella
  15     2320   49.9    954  2  Q9EUE1               Q9eue1 pasteurella
  16     2307   49.6    953  2  Q9EV30               Q9ev30 pasteurella
  17     2286   49.2    955  2  Q9EV22               Q9ev22 pasteurella
  18     2274   48.9    955  2  Q9EUD4               Q9eud4 pasteurella
  19   2229.5   48.0    956  2  Q93NP1               Q93np1 actinobacil
  20     2171   46.7    987  2  Q8KWZ6               Q8kwz6 actinobacil
  21     2155   46.4    987  2  Q8KWZ9               Q8kwz9 actinobacil
  22   2004.5   43.1   1049  2  Q9RCG8               Q9rcg8 pasteurella
  23   1956.5   42.1   1052  2  Q93NP0               Q93np0 actinobacil
  24   1848.5   39.8   1024  16  Q8FE01               Q8fe01 escherichia
  25   1845.5   39.7   1024  2  Q8GA40               Q8ga40 escherichia
  26   1836.5   39.5   1024  2  Q8G9Z4               Q8g9z4 escherichia
  27     1816   39.1   1055  2  Q43892               Q43892 actinobacil
  28     1762   37.9    998  2  O85101               O85101 escherichia
  29     1760   37.9    998  2  Q9LC58               Q9lc58 escherichia
  30     1756   37.8    998  2  P71223               P71223 escherichia
  31     1748   37.6    998  2  Q46716               Q46716 escherichia
  32     1736   37.4   1022  2  Q93NP2               Q93np2 actinobacil
  33     1734   37.3    998  2  Q47461               Q47461 escherichia
  34     1729   37.2    998  2  Q47262               Q47262 escherichia
  35      799   17.2   1706  2  Q9L469               Q9l469 bordetella
  36      482   10.4    233  2  Q51865               Q51865 pasteurella
  37      415    8.9   1636  16  Q9PEL7               Q9pel7 xylella fas
  38      411    8.8   1208  16  Q9PFI9               Q9pfi9 xylella fas
  39      391    8.4   2064  16  Q9PAT8               Q9pat8 xylella fas
  40      383    8.2    208  2  Q51868               Q51868 pasteurella
  41      367    7.9    850  2  Q937V6               Q937v6 bordetella
  42      366    7.9    697  16  Q9JTU8               Q9jtu8 neisseria m
  43      365    7.9    850  2  Q937W0               Q937w0 bordetella
  44    364.5    7.8   1499  16  Q8Y2T5               Q8y2t5 ralstonia s
  45      364    7.8    850  2  Q932Y7               Q932y7 bordetella

                              ALIGNMENTS

RESULT 1
Q93GI2
ID   Q93GI2      PRELIMINARY;      PRT;   927 AA.
AC   Q93GI2;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   RTX toxin.
GN   MBXA.
OS   Moraxella bovis.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC   Moraxellaceae; Moraxella.
OX   NCBI_TaxID=476;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Tifton I;
RX   MEDLINE=21388402; PubMed=11497442;
RA   Angelos J.A., Hess J.F., George L.W.;
RT   "Cloning and characterization of a Moraxella bovis cytotoxin gene.";
RL   Am. J. Vet. Res. 62:1222-1228(2001).
DR   EMBL; AF205359; AAK84651.1; -.
DR   InterPro; IPR002048; EF-hand.
DR   InterPro; IPR001343; Hemlysn_Ca_bind.
DR   InterPro; IPR003995; RtxA.
DR   Pfam; PF00353; hemolysinCabind; 5.
DR   Pfam; PF02382; RTX; 1.
DR   PRINTS; PR00313; CABNDNGRPT.
DR   PRINTS; PR01488; RTXTOXINA.
DR   PROSITE; PS00018; EF_HAND; 1.
DR   PROSITE; PS00330; HEMOLYSIN_CALCIUM; 1.
SQ   SEQUENCE   927 AA;  98845 MW;  F4B703577E10A96D CRC64;

  Query Match             99.9%;  Score 4642;  DB 2;  Length 927;
  Best Local Similarity   99.9%;  Pred. No. 1.6e-199;
  Matches  926;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 MSNINVIKSNIQAGLNSTKSGLKNLYLAIPKDYDPQKGGTLNDFIKAADELGIARLAEEP 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MSNINVIKSNIQAGLNSTKSGLKNLYLAIPKDYDPQKGGTLNDFIKAADELGIARLAEEP 60
```

```
Qy         61 NHTETAKKSVDTVNQFLSLTQTGIAISATKLEKFLQKHSTNKLAKGLDSVENIDRKLGKA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 NHTETAKKSVDTVNQFLSLTQTGIAISATKLEKFLQKHSTNKLAKGLDSVENIDRKLGKA 120

Qy        121 SNVLSTLSSFLGTALAGIELDSLIKKGDAAPDALAKASIDLINEIIGNLSQSTQTIEAFS 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 SNVLSTLSSFLGTALAGIELDSLIKKGDAAPDALAKASIDLINEIIGNLSQSTQTIEAFS 180

Qy        181 SQLAKLGSTISQAKGFSNIGNKLQNLNFSKTNLGLEIITGLLSGISAGFALADKNASTGK 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 SQLAKLGSTISQAKGFSNIGNKLQNLNFSKTNLGLEIITGLLSGISAGFALADKNASTGK 240

Qy        241 KVAAGFELSNQVIGNVTKAISSYVLAQRVAAGLSTTGAVAALITSSIMLAISPLAFMNAA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 KVAAGFELSNQVIGNVTKAISSYVLAQRVAAGLSTTGAVAALITSSIMLAISPLAFMNAA 300

Qy        301 DKFNHANALDEFAKQFRKFGYDGDHLLAEYQRGVGTIEASLTTISTALGAVSAGVSAAAV 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 DKFNHANALDEFAKQFRKFGYDGDHLLAEYQRGVGTIEASLTTISTALGAVSAGVSAAAV 360

Qy        361 GSAVGTPIALLVAGVTGLISGILEASKQAMFESVANRLQGKILEWEKQNGGQNYFDKGYD 420
              ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GSAVGAPIALLVAGVTGLISGILEASKQAMFESVANRLQGKILEWEKQNGGQNYFDKGYD 420

Qy        421 SRYAAYLANNLKFLSELNKELEAERVIAITQQRWDNNIGELAGITKLGERIKSGKAYADA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 SRYAAYLANNLKFLSELNKELEAERVIAITQQRWDNNIGELAGITKLGERIKSGKAYADA 480

Qy        481 FEDGKKVEAGSNITLDAKTGIIDISNSNGKKTQALHFTSPLLTAGTESRERLTNGKYSYI 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 FEDGKKVEAGSNITLDAKTGIIDISNSNGKKTQALHFTSPLLTAGTESRERLTNGKYSYI 540

Qy        541 NKLKFGRVKNWQVTDGEASSKLDFSKVIQRVAETEGTDEIGLIVNAKAGNDDIFVGQGKM 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 NKLKFGRVKNWQVTDGEASSKLDFSKVIQRVAETEGTDEIGLIVNAKAGNDDIFVGQGKM 600

Qy        601 NIDGGDGHDRVFYSKDGGFGNITVDGTSATEAGSYTVNRKVARGDIYHEVVKRQETKVGK 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 NIDGGDGHDRVFYSKDGGFGNITVDGTSATEAGSYTVNRKVARGDIYHEVVKRQETKVGK 660

Qy        661 RTETIQYRDYELRKVGYGYQSTDNLKSVEEVIGSQFNDVFKGSKFNDIFHSGEGDDLLDG 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 RTETIQYRDYELRKVGYGYQSTDNLKSVEEVIGSQFNDVFKGSKFNDIFHSGEGDDLLDG 720

Qy        721 GAGDDRLFGGKGNDRLSGDEGDDLLDGGSGDDVLNGGAGNDVYIFRKGDGNDTLYDGTGN 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GAGDDRLFGGKGNDRLSGDEGDDLLDGGSGDDVLNGGAGNDVYIFRKGDGNDTLYDGTGN 780

Qy        781 DKLAFADANISDIMIERTKEGIIVKRNDHSGSINIPRWYITSNLQNYQSNKTDHKIEQLI 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 DKLAFADANISDIMIERTKEGIIVKRNDHSGSINIPRWYITSNLQNYQSNKTDHKIEQLI 840

Qy        841 GKDGSYITSDQIDKILQDKKDGTVITSQELKKLADENKSQKLSASDIASSLNKLVGSMAL 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GKDGSYITSDQIDKILQDKKDGTVITSQELKKLADENKSQKLSASDIASSLNKLVGSMAL 900

Qy        901 FGTANSVSSNALQPITQPTQGILAPSV 927
              |||||||||||||||||||||||||||
Db        901 FGTANSVSSNALQPITQPTQGILAPSV 927
```

```
RESULT 2
Q9EV24
ID   Q9EV24      PRELIMINARY;      PRT;   946 AA.
AC   Q9EV24;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Leukotoxin.
GN   LKTA.
OS   Mannheimia glucosida.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC   Pasteurellaceae; Mannheimia.
OX   NCBI_TaxID=85401;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PH574;
RX   MEDLINE=21101823; PubMed=11157953;
RA   Davies R.L., Whittam T.S., Selander R.K.;
RT   "Sequence Diversity and Molecular Evolution of the Leukotoxin (lktA)
RT   Gene in Bovine and Ovine Strains of Mannheimia (Pasteurella)
RT   haemolytica.";
RL   J. Bacteriol. 183:1394-1404(2001).
DR   EMBL; AF314521; AAG40305.1; -.
DR   InterPro; IPR001753; EnCoA_hydrtse.
DR   InterPro; IPR001343; Hemlysn_Ca_bind.
DR   InterPro; IPR003995; RtxA.
DR   Pfam; PF00353; hemolysinCabind; 5.
DR   Pfam; PF02382; RTX; 1.
DR   PRINTS; PR00313; CABNDNGRPT.
DR   PRINTS; PR01488; RTXTOXINA.
DR   PROSITE; PS00166; ENOYL_COA_HYDRATASE; 1.
SQ   SEQUENCE   946 AA;  101480 MW;  25C077858BDC76C4 CRC64;

  Query Match             50.8%;  Score 2358.5;  DB 2;  Length 946;
  Best Local Similarity   49.9%;  Pred. No. 1.9e-97;
  Matches  471;  Conservative  182;  Mismatches  258;  Indels   33;  Gaps   15;

Qy          1 MSNINV-IKSN---IQAGLNST----KSGLKNLYLAIPKD--YDPQKGGTLNDFIKAADE 50
Db          5 LTNISTNLKSSWLTAKSGLNRTGQSLKTGAKKIILYIPKDYQYDTEKGNGLQDLVKAAEE 64

Qy         51 LGIARLAEEPNHTETAKKSVDTVNQFLSLTQTGIAISATKLEKFLQKHSTNKLAKGLDSV 110
Db         65 LGIEVQKEEGNDIAKAQTSLGTIQNVLGLTERGIVLSAPQLDKLLQK---TKVGQAIGSA 121

Qy        111 ENIDRKLGKASNVLSTLSSFLGTALAGIELDSLIKKGDAAPDALAKASIDLINEIIGNLS 170
Db        122 ENLTKGFSNAKTVLSGIQSILGSVLAGMDLDEALQK-NSNELTLAKAGLELTNSLIENIA 180

Qy        171 QSTQTIEAFSSQLAKLGSTISQAKGFSNIGNKLQNLN-FSKTNLGLEIITGLLSGISAGF 229
Db        181 NSVKTLDAFGDQINQLGSKLQNVKGLSSLGDKLKGLSGFDKTSLGLDVVSGLLSGATAAL 240

Qy        230 ALADKNASTGKKVAAGFELSNQVIGNVTKAISSYVLAQRVAAGLSTTGAVAALITSSIML 289
Db        241 VLADKNASTSRKVGAGFELANQVVGNITKAVSSYILAQRVAAGLSSTGPVAALIASTVSL 300

Qy        290 AISPLAFMNAADKFNHANALDEFAKQFRKFGYDGDHLLAEYQRGVGTIEASLTTISTALG 349
Db        301 AISPLAFAGIADKFNHAKSLESYAERFKKLGYDGDNLLAEYQRGTGTIDASVTAINTALA 360

Qy        350 AVSAGVSAAAVGSAVGTPIALLVAGVTGLISGILEASKQAMFESVANRLQGKILEWEKQN 409
Db        361 AIAGGVSAAAAGSVIASPIALLVSGITGVISTILQYSKQAMFEHVANKIHNKIVEWEKNN 420

Qy        410 GGQNYFDKGYDSRYAAYLANNLKFLSELNKELEAERVIAITQQRWDNNIGELAGITKLGE 469
Db        421 HGKNYFENGYDARYLANLQDNMKFLLNLNKELQAERVIAITQQQWDNNIGDLAGISRLGE 480

Qy        470 RIKSGKAYADAFEDGKKVEAGSNITLDAKTGIIDISNSNGKKTQALHFTSPLLTAGTESR 529
Db        481 KVLSGKAYVDAFEEGKHLKADKLVQLDSANGIIDVSNSGKAKTQHILFRTPLLTPGTEHR 540

Qy        530 ERLTNGKYSYINKLKFGRVKNWQVTDGEASSKLDFSKVIQRV-------AETEGTDEIGL 582
Db        541 ERVQTGKYEYITKLNINRVDSWKITDGAASSTFDLTNVVQRIGIELDNAGNVTKTKETKI 600

Qy        583 IVNAKAGNDDIFVGQGKMNIDGGDGHDRVFYSKDGGFGNITVDGTSATEAGSYTVNRKVA 642
Db        601 VAKLGAGDDNVFVGSGTTEIDGGEGYDRVHYSR-GNYGALTIDATKETEQGSYTVNRFVE 659

Qy        643 RGDIYHEVVKRQETKVGKRTETIQYRDYELRKVGYGYQSTDNLKSVEEVIGSQFNDVFKG 702
Db        660 TGKALHEVTSTHTALVGNREEKIEYR-HSNNQHHAGYYTKDTLKAVEEIIGTSHNDIFKG 718
```

```
Qy        703 SKFNDIFHSGEGDDLLDGGAGDDRLFGGKGNDRLSGDEGDDLLDGGSGDDVLNGGAGNDV 762
              |||||  |:  :|  |  :||   |:||||||||:|  :  |  |||  :|||  |:|:|:||  |:|:
Db        719 SKFNDAFNGGDGVDTIDGNDGNDRLFGGKGDDIIDGGNGDDFIDGGKGNDLLHGGKGDDI 778

Qy        763 YIFRKGDGNDTLYDGTGNDKLAFADANISDIMIERTKEGIIVKRNDHSGSINIPRWY--- 819
              :: |:||||| :  |  ||||:|:|:|:  |:  |:  |  :::  |        :||  |:
Db        779 FVHRQGDGNDIITDSDGNDKLSFSDSNLKDLTFEKVKHNLVI-TNSRKEKVTIQDWFREA 837

Qy        820 -ITSNLQNYQSNKTDHKIEQLIGKDGSYITSDQIDKILQDKKDGTVITSQELKKLADENK 878
                      ::||::  |  |  |||::||::|  |||  |:|  ::    |      ||  ||  |:  |   :
Db        838 DFAKEVRNYKATK-DEKIEEIIGQNGERITSKQVDDLI--AKGNGKITQDELSKVVDNYE 894

Qy        879 SQKLSASDIASSLNKLVGSMALFGTANSVSSNALQPITQPTQGI 922
                |  |   ::  :||:||:  |  :  |  ::|     :    :  |  :     |  :
Db        895 LLKHS-KNVTNSLDKLISSASAFTSSNDSRNVLVAPTSMLDQSL 937
```

RESULT 3
Q9EV23
ID   Q9EV23   PRELIMINARY;   PRT;   953 AA.
AC   Q9EV23;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Leukotoxin.
GN   LKTA.
OS   Mannheimia glucosida.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC   Pasteurellaceae; Mannheimia.
OX   NCBI_TaxID=85401;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PH290;
RX   MEDLINE=21101823; PubMed=11157953;
RA   Davies R.L., Whittam T.S., Selander R.K.;
RT   "Sequence Diversity and Molecular Evolution of the Leukotoxin (lktA)
RT   Gene in Bovine and Ovine Strains of Mannheimia (Pasteurella)
RT   haemolytica.";
RL   J. Bacteriol. 183:1394-1404(2001).
DR   EMBL; AF314522; AAG40306.1; -.
DR   InterPro; IPR001753; EnCoA_hydrtse.
DR   InterPro; IPR001343; Hemlysn_Ca_bind.
DR   InterPro; IPR003995; RtxA.
DR   Pfam; PF00353; hemolysinCabind; 5.
DR   Pfam; PF02382; RTX; 1.
DR   PRINTS; PR00313; CABNDNGRPT.
DR   PRINTS; PR01488; RTXTOXINA.
DR   PROSITE; PS00166; ENOYL_COA_HYDRATASE; 1.
SQ   SEQUENCE   953 AA; 102150 MW; D99C36DA595B1624 CRC64;

  Query Match   50.7%; Score 2356; DB 2; Length 953;
  Best Local Similarity   49.5%; Pred. No. 2.5e-97;
  Matches 471; Conservative 182; Mismatches 258; Indels   40; Gaps   15;

```
Qy          1 MSNINV-IKSN---IQAGLNST-----------KSGLKNLYLAIPKD--YDPQKGGTLND 43
                ::||:  :||:    ::|||  |              |:|  |  :  |  ||||   ||  :||   |  |
Db          5 LTNISTNLKSSWLTAKSGLNRTGQSLAKAGQSLKTGAKKIILYIPKDYQYDTEKGNGLQD 64

Qy         44 FIKAADELGIARLAEEPNHTETAKKSVDTVNQFLSLTQTGIAISATKLEKFLQKHSTNKL 103
               :|||:||||       ||  |      |:  |:  |:    |  ||:  ||  :||  :|:|  |||       |:
Db         65 LVKAAEELGIEVQKEEGNDIAKAQTSLGTIQNVLGLTERGIVLSAPQLDKLLQK---TKV 121

Qy        104 AKGLDSVENIDRKLGKASNVLSTLSSFLGTALAGIELDSLIKKGDAAPDALAKASIDLIN 163
               :  :  |  ||:  :     |  |||  :  |  ||:  |||:|  ::|  ::       ||||  ::|  |
Db        122 GQAIGSAENLTKGFSNAKTVLSGIQSILGSVLAGMDLDEALQK-NSNELTLAKAGLELTN 180

Qy        164 EIIGNLSQSTQTIEAFSSQLAKLGSTISQAKGFSNIGNKLQNLN-FSKTNLGLEIITGLL 222
               :|  |::  |  :|::||   |:  :|||  :      ||  |::|:|:||:  |:  |  ||:|||::::|||
Db        181 SLIENIANSVKTLDAFGDQINQLGSKLQNVKGLSSLGDKLKGLSGFDKTSLGLDVVSGLL 240

Qy        223 SGISAGFALADKNASTGKKVAAGFELSNQVIGNVTKAISSYVLAQRVAAGLSTTGAVAAL 282
               ||  :|    ||||||||  :||  |||||:|||:||:|||:|||:|||||||||||:||  ||||
Db        241 SGATAALVLADKNASTSRKVGAGFELANQVVGNITKAVSSYILAQRVAAGLSSTGPVAAL 300

Qy        283 ITSSIMLAISPLAFMNAADKFNHANALDEFAKQFRKFGYDGDHLLAEYQRGVGTIEASLT 342
               |  |::  ||||||||      ||||||  :|:  :|::|:|  |||||:||||||||  |||:||:|
Db        301 IASTVSLAISPLAFAGIADKFNHAKSLESYAERFKKLGYDGDNLLAEYQRGTGTIDASVT 360

Qy        343 TISTALGAVSAGVSAAAVGSAVGTPIALLVAGVTGLISGILEASKQAMFESVANRLQGKI 402
               |:|||  |::  ||||||  ||  :  :||||||:|:||:||  ||:  |||||||  |||::  ||
Db        361 AINTALAAIAGGVSAAAAGSVIASPIALLVSGITGVISTILQYSKQAMFEHVANKIHNKI 420

Qy        403 LEWEKQNGGQNYFDKGYDSRYAAYLANNLKFLSELNKELEAERVIAITQQRWDNNIGELA 462
              :||||  |   |:|||:  |||:||  |  |  :|:|||   |||||:||||||||||:||||||:||
Db        421 VEWEKNNHGKNYFENGYDARYLANLQDNMKFLLNLNKELQAERVIAITQQQWDNNIGDLA 480

Qy        463 GITKLGERIKSGKAYADAFEDGKKVEAGSNITLDAKTGIIDISNSNGKKTQALHFTSPLL 522
              ||::|||::  |||||  ||||:||  ::|      :  ||:   ||||:|||    ||  :  |  :|||
Db        481 GISRLGEKVLSGKAYVDAFEEGKHLKADKLVQLDSANGIIDVSNSGKAKTQHILFRTPLL 540

Qy        523 TAGTESRERLTNGKYSYINKLKFGRVKNWQVTDGEASSKLDFSKVIQRV-------AETE 575
               |  |||  |||:   |||  ||  ||     ||  :|::|||  |||    |   :  |:||:
Db        541 TPGTEHRERIQTGKYEYITKLNINRVDSWKITDGAASSTFDLTNVVQRIGIELDNAGNVT 600

Qy        576 GTDEIGLIVNAKAGNDDIFVGQGKMNIDGGDGHDRVFYSKDGGFGNITVDGTSATEAGSY 635
               |  |    ::       ||:|::|||  |      ||||:|:|||  |||:  |  :|  :|:|  |      ||  |||
Db        601 KTKETKIVAKLGAGDDNVFVGSGTTEIDGGEGYDRVHYSR-GNYGALTIDATKETEQGSY 659

Qy        636 TVNRKVARGDIYHEVVKRQETKVGKRTETIQYRDYELRKVGYGYQSTDNLKSVEEVIGSQ 695
              ||||  |    |      |||               ||  |  |  |  |:||  :     :      ||  :  |  ||:|||:||:
Db        660 TVNRFVETGKALHEVTSTHTALVGNREEKIEYR-HSNNQHAGYYTKDTLKAVEEIIGTS 718

Qy        696 FNDVFKGSKFNDIFHSGEGDDLLDGGAGDDRLFGGKGNDRLSGDEGDDLLDGGSGDDVLN 755
               ||:|||||||||  |:   |:  |  |    :||     |:||||||||||:|   :   |    |||  :||  |:|:|:|
Db        719 HNDIFKGSKFNDAFNGGDGVDTIDGNDGNDRLFGGKGDDIIDGGNGDDFIDGGKGNDLLH 778

Qy        756 GGAGNDVYIFRKGDGNDTLYDGTGNDKLAFADANISDIMIERTKEGIIVKRNDHSGSINI 815
               ||  |:|:::   |:||||| :  |      ||||:|:|:|:|:  |:   |:   |:    :::  |        :  |
Db        779 GGKGDDIFVHRQGDGNDIITDSDGNDKLSFSDSNLKDLTFEKVKHNLVI-TNSRKEKVTI 837

Qy        816 PRWY----ITSNLQNYQSNKTDHKIEQLIGKDGSYITSDQIDKILQDKKDGTVITSQELK 871
                   |:                ::||::  |  |  |||::||::|  |||  |:|  ::    |      ||  ||
Db        838 QDWFREADFAKEVRNYKATK-DEKIEEIIGQNGERITSKQVDDLI--AKGNGKITQDELS 894

Qy        872 KLADENKSQKLSASDIASSLNKLVGSMALFGTANSVSSNALQPITQPTQGI 922
               |:  |    :    |  |     ::  :||:||:  |  :  |  ::|     :    :  |  :     |  :
Db        895 KVVDNYELLKHS-KNVTNSLDKLISSASAFTSSNDSRNVLVAPTSMLDQSL 944
```

RESULT 4
Q9ETX2
ID   Q9ETX2   PRELIMINARY;   PRT;   953 AA.
AC   Q9ETX2;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Leukotoxin.
GN   LKTA.
OS   Mannheimia glucosida.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC   Pasteurellaceae; Mannheimia.
OX   NCBI_TaxID=85401;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PH498, and PH344;
RA   Davies R.L., Whittam T.S., Selander R.K.;
RT   "Sequence diversity and molecular evolution of the leukotoxin (lktA)
RT   gene in bovine and ovine strains of Mannheimia (Pasteurella)
RT   haemolytica.";
RL   J. Bacteriol. 0:0-0(2001).
DR   EMBL; AF314518; AAG40302.1; -.
DR   EMBL; AF314517; AAG40301.1; -.
DR   InterPro; IPR001753; EnCoA_hydrtse.
DR   InterPro; IPR001343; Hemlysn_Ca_bind.
DR   InterPro; IPR003995; RtxA.

```
DR   Pfam; PF00353; hemolysinCabind; 5.
DR   Pfam; PF02382; RTX; 1.
DR   PRINTS; PR00313; CABNDNGRPT.
DR   PRINTS; PR01488; RTXTOXINA.
DR   PROSITE; PS00166; ENOYL_COA_HYDRATASE; 1.
SQ   SEQUENCE   953 AA;  102135 MW;  70DB354157F5881E CRC64;

  Query Match             50.7%;  Score 2355;  DB 2;  Length 953;
  Best Local Similarity   49.5%;  Pred. No. 2.8e-97;
  Matches  471;  Conservative  182;  Mismatches  258;  Indels   40;  Gaps   15;

Qy          1 MSNINV-IKSN---IQAGLNST-----------KSGLKNLYLAIPKD--YDPQKGGTLND 43
              ::||:  :||:     ::|||  |                |:|  |  :  | ||||   || :|| |  |
Db          5 LTNISTNLKSSWLTAKSGLNRTGQSLAKAGQSLKTGAKKIILYIPKDYQYDTEKGNGLQD 64

Qy         44 FIKAADELGIARLAEEPNHTETAKKSVDTVNQFLSLTQTGIAISATKLEKFLQKHSTNKL 103
               :|||:||||     || |       |: |: |:    | ||: || :|| :|:| |||     |:
Db         65 LVKAAEELGIEVQKEEGNDIAKAQTSLGTIQNVLGLTERGIVLSAPQLDKLLQK---TKV 121

Qy        104 AKGLDSVENIDRKLGKASNVLSTLSSFLGTALAGIELDSLIKKGDAAPDALAKASIDLIN 163
               : : | ||: :       |  |||  : | ||: |||::||   ::| ::      |||| ::| |
Db        122 GQAIGSSAENLTKGFSNAKTVLSGIQSILGSVLAGMDLDEALQK-NSNELTLAKAGLELTN 180

Qy        164 EIIGNLSQSTQTIEAFSSQLAKLGSTISQAKGFSNIGNKLQNLN-FSKTNLGLEIITGLL 222
               :| |:: | :|::||  |: :||| :     || |::|:||: |:  | ||:|||::::|||
Db        181 SLIENIANSVKTLDAFGDQINQLGSKLQNVKGLSSLGDKLKGLSGFDKTSLGLDVVSGLL 240

Qy        223 SGISAGFALADKNASTGKKVAAGFELSNQVIGNVTKAISSYVLAQRVAAGLSTTGAVAAL 282
              || :|     ||||||| :||  ||||:|||:||:||:|||:||:||||||||||:||  ||||
Db        241 SGATAALVLADKNASTSRKVGAGFELANQVVGNITKAVSSYILAQRVAAGLSSTGPVAAL 300

Qy        283 ITSSIMLAISPLAFMNAADKFNHANALDEFAKQFRKFGYDGDHLLAEYQRGVGTIEASLT 342
               | |:: |||||||    ||||||  :|:  :|::|:| |||||:||||||||  |||:|:|:|
Db        301 IASTVSLAISPLAFAGIADKFNHAKSLESYAERFKKLGYDGDNLLAEYQRGTGTIDASVT 360

Qy        343 TISTALGAVSAGVSAAAVGSAVGTPIALLVAGVTGLISGILEASKQAMFESVANRLQGKI 402
               |:||| |:: |||||| ||  : :||||||:|:||:|| ||: ||||||| |||::   ||
Db        361 AINTALAAIAGGVSAAAAGSVIASPIALLVSGITGVISTILQYSKQAMFEHVANKIHNKI 420

Qy        403 LEWEKQNGGQNYFDKGYDSRYAAYLANNLKFLSELNKELEAERVIAITQQRWDNNIGELA 462
              :||||  |  |:|||:  |||:||  |  |  :|:|||   |||||:||||||||||:||||||:||
Db        421 VEWEKNNHGKNYFENGYDARYLANLQDNMKFLLNLNKELQAERVIAITQQQWDNNIGDLA 480

Qy        463 GITKLGERIKSGKAYADAFEDGKKVEAGSNITLDAKTGIIDISNSNGKKTQALHFTSPLL 522
              ||:||||:: ||||| ||||:|  ||||  :|:    ||:  ||||||:||  ||| : |  :||
Db        481 GISRLGEKVLSGKAYVDAFEEGKHLKADKLVQLDSANGIIDVSNSGKAKTQHILFRTPLL 540

Qy        523 TAGTESRERLTNGKYSYINKLKFGRVKNWQVTDGEASSKLDFSKVIQRV--------AETE 575
               | ||| |||:   ||| || ||    ||  |:|::||| |||     |  :  |:||:
Db        541 TPGTEHRERVQTGKYEYITKLNINRVDSWKITDGAASSTFDLTNVVQRIGIELDNAGNVT 600

Qy        576 GTDEIGLIVNAKAGNDDIFVGQGKMNIDGGDGHDRVFYSKDGGFGNITVDGTSATEAGSY 635
               |  |    ::       ||:|::|||  |      ||||:|:|||  ||:  |  :|  :|:|  |    || |||
Db        601 KTKETKIVAKLGAGDDNVFVGSGTTEIDGGEGYDRVHYSR-GNYGALTIDATKETEQGSY 659

Qy        636 TVNRKVARGDIYHEVVKRQETKVGKRTETIQYRDYELRKVGYGYQSTDNLKSVEEVIGSQ 695
              |||| |   |    |||             || |  |  |:||  :      :    ||  :  |  ||:|||:||:
Db        660 TVNRFVETGKALHEVTSTHTALVGNREEKIEYR-HSNNQHHAGYYTKDTLKAVEEIIGTS 718

Qy        696 FNDVFKGSKFNDIFHSGEGDDLLDGGAGDDRLFGGKGNDRLSGDEGDDLLDGGSGDDVLN 755
               |:|||||||||  |: |:  |  :|| |:||||||||:|  |    |||  :|||  |:|:|:
Db        719 HNDIFKGSKFNDAFNGGDGVDTIDGNDGNDRLFGGKGDDIIDGGNGDDFIDGGKGNDLLH 778

Qy        756 GGAGNDVYIFRKGDGNDTLYDGTGNDKLAFADANISDIMIERTKEGIIVKRNDHSGSINI 815
              || |:|::: |:||||| :  |||||:|:|:|: |:   |:  |  |  :::    |     |
Db        779 GGKGDDIFVHRQGDGNDIITDSDGNDKLSFSDSNLKDLTFEKVKHNLVI-TNSRKEKVTI 837

Qy        816 PRWY----ITSNLQNYQSNKTDHKIEQLIGKDGSYITSDQIDKILQDKKDGTVITSQELK 871
               |:           ::||::  | | |||::||::|   |||  |:|  ::       ||    ||
Db        838 QDWFREADFAKEVRNYKATK-DEKIEEIIGQNGERITSKQVDDLI--AKGNGKITQDELS 894

Qy        872 KLADENKSQKLSASDIASSLNKLVGSMALFGTANSVSSNALQPITQPTQGI 922
              |: |   :   | |   :: :||:||: | : | ::|     :   : | :    | :
Db        895 KVVDNYELLKHS-KNVTNSLDKLISSASAFTSSNDSRNVLVAPTSMLDQSL 944

RESULT 5
Q9EV25
ID   Q9EV25      PRELIMINARY;      PRT;   953 AA.
AC   Q9EV25;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Leukotoxin.
GN   LKTA.
OS   Mannheimia glucosida.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC   Pasteurellaceae; Mannheimia.
OX   NCBI_TaxID=85401;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PH496;
RX   MEDLINE=21101823; PubMed=11157953;
RA   Davies R.L., Whittam T.S., Selander R.K.;
RT   "Sequence Diversity and Molecular Evolution of the Leukotoxin (lktA)
RT   Gene in Bovine and Ovine Strains of Mannheimia (Pasteurella)
RT   haemolytica.";
RL   J. Bacteriol. 183:1394-1404(2001).
DR   EMBL; AF314520; AAG40304.1; -.
DR   InterPro; IPR001753; EnCoA_hydrtse.
DR   InterPro; IPR001343; Hemlysn_Ca_bind.
DR   InterPro; IPR003995; RtxA.
DR   Pfam; PF00353; hemolysinCabind; 5.
DR   Pfam; PF02382; RTX; 1.
DR   PRINTS; PR00313; CABNDNGRPT.
DR   PRINTS; PR01488; RTXTOXINA.
DR   PROSITE; PS00166; ENOYL_COA_HYDRATASE; 1.
SQ   SEQUENCE   953 AA;  102138 MW;  C41D9EBC1D799951 CRC64;

  Query Match             50.7%;  Score 2355;  DB 2;  Length 953;
  Best Local Similarity   49.5%;  Pred. No. 2.8e-97;
  Matches  471;  Conservative  182;  Mismatches  258;  Indels   40;  Gaps   15;

Qy          1 MSNINV-IKSN---IQAGLNST-----------KSGLKNLYLAIPKD--YDPQKGGTLND 43
              ::||:  :||:     ::|||  |                |:|  |  :  | ||||   || :|| |  |
Db          5 LTNISTNLKSSWLTAKSGLNRTGQSLAKAGQSLKTGAKKIILYIPKDYQYDTEKGNGLQD 64

Qy         44 FIKAADELGIARLAEEPNHTETAKKSVDTVNQFLSLTQTGIAISATKLEKFLQKHSTNKL 103
               :|||:||||     || |       |: |: |:    | ||: || :|| :|:| |||     |:
Db         65 LVKAAEELGIEVQKEEGNDIAKAQTSLGTIQNVLGLTERGIVLSAPQLDKLLQK---TKV 121

Qy        104 AKGLDSVENIDRKLGKASNVLSTLSSFLGTALAGIELDSLIKKGDAAPDALAKASIDLIN 163
               : : | ||: :       |  |||  : | ||: |||::||   ::| ::      |||| ::| |
Db        122 GQAIGSSAENLTKGFSNAKTVLSGIQSILGSVLAGMDLDEALQK-NSNELTLAKAGLELTN 180

Qy        164 EIIGNLSQSTQTIEAFSSQLAKLGSTISQAKGFSNIGNKLQNLN-FSKTNLGLEIITGLL 222
               :| |:: | :|::||  |: :||| :     || |::|:||: |:  | ||:|||::::|||
Db        181 SLIENIANSVKTLDAFGDQINQLGSKLQNVKGLSSLGDKLKGLSGFDKTSLGLDVVSGLL 240

Qy        223 SGISAGFALADKNASTGKKVAAGFELSNQVIGNVTKAISSYVLAQRVAAGLSTTGAVAAL 282
              || :|     ||||||| :||  ||||:|||:||:||:|||:||:||||||||||:||  ||||
Db        241 SGATAALVLADKNASTSRKVGAGFELANQVVGNITKAVSSYILAQRVAAGLSSTGPVAAL 300

Qy        283 ITSSIMLAISPLAFMNAADKFNHANALDEFAKQFRKFGYDGDHLLAEYQRGVGTIEASLT 342
               | |:: |||||||    ||||||  :|:  :|::|:| |||||:||||||||  |||:|:|:|
Db        301 IASTVSLAISPLAFAGIADKFNHAKSLESYAERFKKLGYDGDNLLAEYQRGTGTIDASVT 360

Qy        343 TISTALGAVSAGVSAAAVGSAVGTPIALLVAGVTGLISGILEASKQAMFESVANRLQGKI 402
               |:||| |:: |||||| ||  : :||||||:|:||:|| ||: ||||||| |||::   ||
Db        361 AINTALAAIAGGVSAAAAGSVIASPIALLVSGITGVISTILQYSKQAMFEHVANKIHNKI 420

Qy        403 LEWEKQNGGQNYFDKGYDSRYAAYLANNLKFLSELNKELEAERVIAITQQRWDNNIGELA 462
              :||||  |  |:|||:  |||:||  |  |  :|:|||   |||||:||||||||||:||||||:||
Db        421 VEWEKNNHGKNYFENGYDARYLANLQDNMKFLLNLNKELQAERVIAITQQQWDNNIGDLA 480

Qy        463 GITKLGERIKSGKAYADAFEDGKKVEAGSNITLDAKTGIIDISNSNGKKTQALHFTSPLL 522
```

```
             ||::|||:: ||||| ||||:|| ::|    : ||:  ||||:|||    ||| : | :|||
Db       481 GISRLGEKVLSGKAYVDAFEEGKHLKADKLVQLDSANGIIDVSNSGKAKTQHILFRTPLL 540

Qy       523 TAGTESRERLTNGKYSYINKLKFGRVKNWQVTDGEASSKLDFSKVIQRV-------AETE 575
             | ||| |||:  ||| || ||   || :|::||| |||   | : |:||:
Db       541 TPGTEHRERVQTGKYEYITKLNINRVDSWKITDGAASSTFDLTNVVQRIGIELDNAGNVT 600

Qy       576 GTDEIGLIVNAKAGNDDIFVGQGKMNIDGGDGHDRVFYSKDGGFGNITVDGTSATEAGSY 635
             | |  ::     ||:|::||| |    ||||:|:||| ||:  | :| :|:| |   || |||
Db       601 KTKETKIVAKLGAGDDNVFVGSGTTEIDGGEGYDRVHYSR-GNYGALTIDATKETEQGSY 659

Qy       636 TVNRKVARGDIYHEVVKRQETKVGKRTETIQYRDYELRKVGYGYQSTDNLKSVEEVIGSQ 695
             |||| |   | |||           ||| |   | |:||  :      || : | ||:|||:||:
Db       660 TVNRFVETGKALHEVTSTHTALVGNREEKIEYR-HSNNQHHAGYYTKDTLKAVEEIIGTS 718

Qy       696 FNDVFKGSKFNDIFHSGEGDDLLDGGAGDDRLFGGKGNDRLSGDEGDDLLDGGSGDDVLN 755
             ||:||||||| |: |:| | :||  |:||||||||:| : |  ||| :||| |:|:|:
Db       719 HNDIFKGSKFNDAFNGGDGVDTIDGNDGNDRLFGGKGDDIIDGGNGDDFIDGGKGNDLLH 778

Qy       756 GGAGNDVYIFRKGDGNDTLYDGTGNDKLAFADANISDIMIERTKEGIIVKRNDHSGSINI 815
             || |:|::: |:||||| : |  ||||:|:|:|: |: |: | ::: |      : |
Db       779 GGKGDDIFVHRQGDGNDIITDSDGNDKLSFSDSNLKDLTFEKVKHNLVI-TNSKKEKVTI 837

Qy       816 PRWY----ITSNLQNYQSNKTDHKIEQLIGKDGSYITSDQIDKILQDKKDGTVITSQELK 871
              |:         ::||:: | | |||::||::| ||| |:| ::     |     || ||
Db       838 QDWFREADFAKEVRNYKATK-DEKIEEIIGQNGERITSKQVDDLI--AKGNGKITQDELS 894

Qy       872 KLADENKSQKLSASDIASSLNKLVGSMALFGTANSVSSNALQPITQPTQGI 922
             |: |  :  | |  :: :||:||: | : | ::|   :  :  | :   | :
Db       895 KVVDNYELLKHS-KNVTNSLDKLISSASAFTSSNDSRNVLVAPTSMLDQSL 944
```

RESULT 6
Q9EV26
ID   Q9EV26   PRELIMINARY;   PRT;   953 AA.
AC   Q9EV26;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Leukotoxin.
GN   LKTA.
OS   Mannheimia glucosida.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC   Pasteurellaceae; Mannheimia.
OX   NCBI_TaxID=85401;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PH240;
RX   MEDLINE=21101823; PubMed=11157953;
RA   Davies R.L., Whittam T.S., Selander R.K.;
RT   "Sequence Diversity and Molecular Evolution of the Leukotoxin (lktA)
RT   Gene in Bovine and Ovine Strains of Mannheimia (Pasteurella)
RT   haemolytica.";
RL   J. Bacteriol. 183:1394-1404(2001).
DR   EMBL; AF314519; AAG40303.1; -.
DR   InterPro; IPR001753; EnCoA_hydrtse.
DR   InterPro; IPR001343; Hemlysn_Ca_bind.
DR   InterPro; IPR003995; RtxA.
DR   Pfam; PF00353; hemolysinCabind; 5.
DR   Pfam; PF02382; RTX; 1.
DR   PRINTS; PR00313; CABNDNGRPT.
DR   PRINTS; PR01488; RTXTOXINA.
DR   PROSITE; PS00166; ENOYL_COA_HYDRATASE; 1.
SQ   SEQUENCE   953 AA; 102236 MW; AC5874B3B75D1C98 CRC64;

Query Match 50.5%; Score 2345; DB 2; Length 953;
Best Local Similarity 48.9%; Pred. No. 7.9e-97;
Matches 468; Conservative 186; Mismatches 250; Indels 54; Gaps 16;

```
Qy         1 MSNINV-IKSN---IQAGLNST----------KSGLKNLYLAIPKD--YDPQKGGTLND 43
              ::||:  :||:     ::||| |          |:| |  :  | |||| ||:|| |
Db         5 LTNISTNLKSSWLTAKSGLNRTGQSLAKAGQSLKTGAKKIILYIPKDYQYDTEKGNGLQD 64

Qy        44 FIKAADELGIARLAEEPNHTETAKKSVDTVNQFLSLTQTGIAISATKLEKFLQKHSTNKL 103
              :|||:||||     || |       |: |: |:    | ||: || :|| :|:| |||    |:
Db        65 LVKAAEELGIEVQKEEGNDIAKAQTSLGTIQNVLGLTERGIVLSAPQLDKLLQK---TKV 121

Qy       104 AKGLDSVENIDRKLGKASNVLSTLSSFLGTALAGIELDSLIKKGDAAPDALAKASIDLIN 163
             : : | ||: :      | |||  : | ||: |||::||     ::| ::      ||| ::| |
Db       122 GQAIGSAENLTKGFSNAKTVLSGIQSILGSVLAGMDLDEALQK-NSNELTLAKAGLELTN 180

Qy       164 EIIGNLSQSTQTIEAFSSQLAKLGSTISQAKGFSNIGNKLQNLN-FSKTNLGLEIITGLL 222
             :| |:: | :|::|| |: :||| :     || | |::|:|:| |: | ||:|||:::|||
Db       181 SLIENIANSVKTLDAFGDQINQLGSKLQNVKGLSSLGDKLKGLSGFDKTSLGLDVVSGLL 240

Qy       223 SGISAGFALADKNASTGKKVAAGFELSNQVIGNVTKAISSYVLAQRVAAGLSTTGAVAAL 282
             || :|   ||||||| :|| ||||:|||:||:||:|||:|||||||||||:|| ||||
Db       241 SGATAALVLADKNASTSRKVGAGFELANQVVGNITKAVSSYILAQRVAAGLSSTGPVAAL 300

Qy       283 ITSSIMLAISPLAFMNAADKFNHANALDEFAKQFRKFGYDGDHLLAEYQRGVGTIEASLT 342
             | |:: ||||||||   ||||||| :|: :|::|:| |||||:|||||||| |||:||:|
Db       301 IASTVSLAISPLAFAGIADKFNHAKSLESYAERFKKLGYDGDNLLAEYQRGTGTIDASVT 360

Qy       343 TISTALGAVSAGVSAAAVGSAVGTPIALLVAGVTGLISGILEASKQAMFESVANRLQGKI 402
              :|||  |::  ||||| || :  |||||:|:|:|:|:||:|| ||||||||:: ||   ||
Db       361 AINTALAAIAGGVSAAAAGSVIASPIALLVSGITGVISTILQYSKQAMFEHVANKIHNKI 420

Qy       403 LEWEKQNGGQNYFDKGYDSRYAAYLANNLKFLSELNKELEAERVIAITQQRWDNNIGELA 462
              :|||| | |:||: |||  | |  | |:||| |||||||||||||:||||||||:||
Db       421 VEWEKNNHGKNYFENGYDARYLANLQDNMKFLLNLNKELQAERVIAITQQQWDNNIGDLA 480

Qy       463 GITKLGERIKSGKAYADAFEDGKKVEAGSNITLDAKTGIIDISNSNGKKTQALHFTSPLL 522
             ||::|||:: ||||| ||||:||  ::|      : ||:  ||||:|||   ||| : :|||
Db       481 GISRLGEKVLSGKAYVDAFEEGKHLKADKLVQLDSANGIIDVSNSGKAKTQHILFRTPLL 540

Qy       523 TAGTESRERLTNGKYSYINKLKFGRVKNWQVTDGEASSKLDFSKVIQRV-------AETE 575
             | ||| |||:  ||| || ||   || :|::||| |||   | : |:||:
Db       541 TPGTEHRERVQTGKYEYITKLNINRVDSWKITDGAASSTFDLTNVVQRIGIELDNAGNVT 600

Qy       576 GTDEIGLIVNAKAGNDDIFVGQGKMNIDGGDGHDRVFYSKDGGFGNITVDGTSATEAGSY 635
             | |  ::     ||:|::||| |    ||||:|:||| ||:  | :| :|:| |   || |||
Db       601 KTKETKIVAKLGAGDDNVFVGSGTTEIDGGEGYDRVHYSR-GNYGALTIDATKETERGSY 659

Qy       636 TVNRKVARGDIYHEVVKRQETKVGKRTETIQYRDYELRKVGYGYQSTDNLKSVEEVIGSQ 695
             |||| |   | |||           ||| |   | |:||  :      || : | ||:|||:||:
Db       660 TVNRFVETGKALHEVTSTHTALVGNREEKIEYR-HSNNQHHAGYYTKDTLKAIEEIIGTS 718

Qy       696 FNDVFKGSKFNDIFHSGEGDDLLDGGAGDDRLFGGKGNDRLSGDEGDDLLDGGSGDDVLN 755
             ||:||||||| |: |:| | :||  |:||||||||:| : |  ||| :||| |:|:|:
Db       719 HNDIFKGSKFNDAFNGGDGVDTIDGNDGNDRLFGGKGDDIIDGGNGDDFIDGGKGNDLLH 778

Qy       756 GGAGNDVYIFRKGDGNDTLYDGTGNDKLAFADANISDIMIERTKEGIIVKRNDHSGSINI 815
             || |:|::: |:||||| : |  ||||:|:|:|: |: |: | ::: |      : |
Db       779 GGKGDDIFVHRQGDGNDIITDSDGNDKLSFSDSNLKDLTFEKVKHNLVI-TNSRKEKVTI 837

Qy       816 PRWY----ITSNLQNYQSNKTDHKIEQLIGKDGSYITSDQIDKILQ------DKKD-GTV 864
              |:         ::||:: | | |||::||::| ||| |:|::::     || |  |
Db       838 QDWFREADFAKEVRNYKATK-DEKIEEIIGQNGERITSKQVDELIEKGKGKIDKSDLSQV 896

Qy       865 ITSQELKKLADENKSQKLSASDIASSLNKLVGSMALFGTANSVSSNALQPITQPTQGI 922
              : : || |              : :: :||:||: | : | ::|   :  :  | :   | :
Db       897 VDNYELLK----------HSKNVTNSLDKLISSASAFTSSNDSRNVLVAPTSMLDQSL 944
```

RESULT 7
Q9EV32
ID   Q9EV32   PRELIMINARY;   PRT;   953 AA.
AC   Q9EV32;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Leukotoxin.
GN   LKTA.
OS   Pasteurella haemolytica.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC   Pasteurellaceae; Mannheimia.

```
OX   NCBI_TaxID=75985;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PH706;
RX   MEDLINE=21101823; PubMed=11157953;
RA   Davies R.L., Whittam T.S., Selander R.K.;
RT   "Sequence Diversity and Molecular Evolution of the Leukotoxin (lktA)
RT   Gene in Bovine and Ovine Strains of Mannheimia (Pasteurella)
RT   haemolytica.";
RL   J. Bacteriol. 183:1394-1404(2001).
DR   EMBL; AF314509; AAG40293.1; -.
DR   InterPro; IPR001753; EnCoA_hydrtse.
DR   InterPro; IPR001343; Hemlysn_Ca_bind.
DR   InterPro; IPR003995; RtxA.
DR   Pfam; PF00353; hemolysinCabind; 5.
DR   Pfam; PF02382; RTX; 1.
DR   PRINTS; PR00313; CABNDNGRPT.
DR   PRINTS; PR01488; RTXTOXINA.
DR   PROSITE; PS00166; ENOYL_COA_HYDRATASE; 1.
SQ   SEQUENCE   953 AA;  102078 MW;  EF425243C8741EE4 CRC64;

  Query Match             50.3%;  Score 2338.5;  DB 2;  Length 953;
  Best Local Similarity   49.5%;  Pred. No. 1.5e-96;
  Matches  473;  Conservative  183;  Mismatches  255;  Indels   45;  Gaps   18;

Qy          1 MSNINV-IKSN---IQAGLNST-----------KSGLKNLYLAIPKD--YDPQKGGTLND 43
Db          5 LTNISTNLKSSWLTAKSGLNRTGQSLAKAGQSLKTGAKKIILYIPKDYQYDTEKGNGLQD 64

Qy         44 FIKAADELGIARLAEEPNHTETAKKSVDTVNQFLSLTQTGIAISATKLEKFLQKHSTNKL 103
Db         65 LVKAAQELGIEVQKEEGNDIAKAQTSLGTIQNVLGLTERGIVLSAPQLDKLLQK---TKV 121

Qy        104 AKGLDSVENIDRKLGKASNVLSTLSSFLGTALAGIELDSLIKKGDAAPDALAKASIDLIN 163
Db        122 GQAIGSAENLTKGFSNAKTVLSGIQSILGSVLAGMDLDEALQK-NSNELTLAKAGLELTN 180

Qy        164 EIIGNLSQSTQTIEAFSSQLAKLGSTISQAKGFSNIGNKLQNLN-FSKTNLGLEIITGLL 222
Db        181 SLIENIANSVKTLDAFGDQINQLGSKLQNVKGLSSLGDKLKGLSGFDKTSLGLDVVSGLL 240

Qy        223 SGISAGFALADKNASTGKKVAAGFELSNQVIGNVTKAISSYVLAQRVAAGLSTTGAVAAL 282
Db        241 SGATAALVLADKNASTSRKVGAGFELANQVVGNITKAVSSYILAQRVAAGLSSTGPVAAL 300

Qy        283 ITSSIMLAISPLAFMNAADKFNHANALDEFAKQFRKFGYDGDHLLAEYQRGVGTIEASLT 342
Db        301 IASTVSLAISPLAFAGIADKFNHAKSLESYAERFKKLGYDGDNLLAEYQRGTGTIDASVT 360

Qy        343 TISTALGAVSAGVSAAAVGSAVGTPIALLVAGVTGLISGILEASKQAMFESVANRLQGKI 402
Db        361 AINTALAAIAGGVSAAAAGSVIASPIALLVSGITGVISTILQYSKQAMFEHVANKIHNKI 420

Qy        403 LEWEKQNGGQNYFDKGYDSRYAAYLANNLKFLSELNKELEAERVIAITQQRWDNNIGELA 462
Db        421 VEWEKNNGGKNYFENGYDARYLANLQDNMKFLLNLNKELQAERVIAITQQQWDSNIGDLA 480

Qy        463 GITKLGERIKSGKAYADAFEDGKKVEAGSNITLDAKTGIIDISNSNGKKTQALHFTSPLL 522
Db        481 GISRLGEKVLSGKAYVDAFEEGQHLKADKLVQLDSAKGIIDVSNTGEAKTQHILFRTPLL 540

Qy        523 TAGTESRERLTNGKYSYINKLKFGRVKNWQVTDGEASSKLDFSKVIQRV-------AETE 575
Db        541 TPGTEKRERVQTGKYEYITKLHINRVDSWKITDGAASSTFDLTNVVQRIGIELDNAGNVT 600

Qy        576 GTDEIGLIVNAKAGNDDIFVGQGKMNIDGGDGHDRVFYSKDGGFGNITVDGTSATEAGSY 635
Db        601 KTKETKIIAKLGEGDDNVFVGSGTTEIDGGEGYDRVHYSR-GNYGALTIDATKETEQGSY 659

Qy        636 TVNRKVARGDIYHEVVKRQETKVGKRTETIQYRDYELRKVGYGYQSTDNLKSVEEVIGSQ 695
Db        660 TVNRFVESGKALHEVTSTHTALVGNREEKIEYR-HSNNQHHAGYYTKDTLKAVEEIIGTS 718

Qy        696 FNDVFKGSKFNDIFHSGEGDDLLDGGAGDDRLFGGKGNDRLSGDEGDDLLDGGSGDDVLN 755
Db        719 HNDIFKGSKFNDAFNGGDGVDTIDGNDGNDRLFGGKGDDIIDGGNGDDFIDGGKGNDLLH 778

Qy        756 GGAGNDVYIFRKGDGNDTLYDGTGNDKLAFADANISDIMIERTKEGIIVKRNDHSGSINI 815
Db        779 GGKGDDIFVHRQGDGNDSITESEGNDKLSFSDSNLKDLTFEKVNHHLVI-TNTKQEKVTI 837

Qy        816 PRWY----ITSNLQNYQSNKTDHKIEQLIGKDGSYITSDQIDKILQDKKDGTVITSQELK 871
Db        838 QNWFREAEFAKTIQNYVATR-DDKIEEIIGQNGERITSKQVDELIE--KGNGKIAQSELT 894

Qy        872 KLADENKSQKLSASDIASSLNKLVGSMALFGTANSVSSNALQPITQPTQGILAPSV 927
Db        895 KVVDNYQLLKYS-RDASNSLDKLISSASAFTSSND-SRNVL---ASPT-SMLDPSL 944
```

RESULT 8
Q9EV33

```
ID   Q9EV33      PRELIMINARY;      PRT;   953 AA.
AC   Q9EV33;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Leukotoxin.
GN   LKTA.
OS   Pasteurella haemolytica.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC   Pasteurellaceae; Mannheimia.
OX   NCBI_TaxID=75985;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PH66;
RX   MEDLINE=21101823; PubMed=11157953;
RA   Davies R.L., Whittam T.S., Selander R.K.;
RT   "Sequence Diversity and Molecular Evolution of the Leukotoxin (lktA)
RT   Gene in Bovine and Ovine Strains of Mannheimia (Pasteurella)
RT   haemolytica.";
RL   J. Bacteriol. 183:1394-1404(2001).
DR   EMBL; AF314508; AAG40292.1; -.
DR   InterPro; IPR001753; EnCoA_hydrtse.
DR   InterPro; IPR001343; Hemlysn_Ca_bind.
DR   InterPro; IPR003995; RtxA.
DR   Pfam; PF00353; hemolysinCabind; 5.
DR   Pfam; PF02382; RTX; 1.
DR   PRINTS; PR00313; CABNDNGRPT.
DR   PRINTS; PR01488; RTXTOXINA.
DR   PROSITE; PS00166; ENOYL_COA_HYDRATASE; 1.
SQ   SEQUENCE   953 AA;  102132 MW;  4138AB5FAE2843B3 CRC64;

  Query Match             50.3%;  Score 2336.5;  DB 2;  Length 953;
  Best Local Similarity   49.6%;  Pred. No. 1.9e-96;
  Matches  474;  Conservative  181;  Mismatches  256;  Indels   45;  Gaps   18;

Qy          1 MSNINV-IKSN---IQAGLNST-----------KSGLKNLYLAIPKD--YDPQKGGTLND 43
Db          5 LTNISTNLKSSWLTAKSGLNRTGQSLAKAGQSLKTGAKKIILYIPKDYQYDTEKGNGLQD 64

Qy         44 FIKAADELGIARLAEEPNHTETAKKSVDTVNQFLSLTQTGIAISATKLEKFLQKHSTNKL 103
Db         65 LVKAAQELGIEVQKEEGNDIAKAQTSLGTIQNVLGLTERGIVLSAPQLDKLLQK---TKV 121

Qy        104 AKGLDSVENIDRKLGKASNVLSTLSSFLGTALAGIELDSLIKKGDAAPDALAKASIDLIN 163
Db        122 GQAIGSAENLTKGFSNAKTVLSGIQSILGSVLAGMDLDEALQK-NSNELTLAKAGLELTN 180

Qy        164 EIIGNLSQSTQTIEAFSSQLAKLGSTISQAKGFSNIGNKLQNLN-FSKTNLGLEIITGLL 222
Db        181 SLIENIANSVKTLDAFGDQINQLGSKLQNVKGLSSLGDKLKGLSGFDKTSLGLDVVSGLL 240

Qy        223 SGISAGFALADKNASTGKKVAAGFELSNQVIGNVTKAISSYVLAQRVAAGLSTTGAVAAL 282
Db        241 SGATAALVLADKNASTSRKVGAGFELANQVVGNITKAVSSYILAQRVAAGLSSTGPVAAL 300
```

```
Qy        283 ITSSIMLAISPLAFMNAADKFNHANALDEFAKQFRKFGYDGDHLLAEYQRGVGTIEASLT 342
              | |:: |||||||     ||||||  :|: :|::|:| |||||:||||||||  |||:||:|
Db        301 IASTVSLAISPLAFAGIADKFNHAKSLESYAERFKKLGYDGDNLLAEYQRGTGTIDASVT 360

Qy        343 TISTALGAVSAGVSAAAVGSAVGTPIALLVAGVTGLISGILEASKQAMFESVANRLQGKI 402
              |:||| |:: ||||| || | :||||||:|:||:|| ||: |||||| |||::  ||
Db        361 AINTALAAIAGGVSAAAAGSVVASPIALLVSGITGVISTILQYSKQAMFEHVANKIHNKI 420

Qy        403 LEWEKQNGGQNYFDKGYDSRYAAYLANNLKFLSELNKELEAERVIAITQQRWDNNIGELA 462
              :|||| | |:|||: |||:|| | | :|:||| |||||:|||||||||||:|||||||:||
Db        421 VEWEKNNPGKNYFENGYDARYLANLQDNMKFLLNLNKELQAERVIAITQQQWDNNIGDLA 480

Qy        463 GITKLGERIKSGKAYADAFEDGKKVEAGSNITLDAKTGIIDISNSNGKKTQALHFTSPLL 522
              ||::|||:: ||||| ||||:|: ::|    : ||:  ||||:||:   ||| : | :|||
Db        481 GISRLGEKVLSGKAYVDAFEEGQHLKADKLVQLDSAKGIIDVSNTGEAKTQHILFRTPLL 540

Qy        523 TAGTESRERLTNGKYSYINKLKFGRVKNWQVTDGEASSKLDFSKVIQRV-------AETE 575
              | ||| |||:  ||| || ||   || :|::||| |||  | : |:||:
Db        541 TPGTEKRERVQTGKYEYITKLHINRVDSWKITDGAASSTFDLTNVVQRIGIELDNAGNVT 600

Qy        576 GTDEIGLIVNAKAGNDDIFVGQGKMNIDGGDGHDRVFYSKDGGFGNITVDGTSATEAGSY 635
              | |  :|   |:|::|||  ||||:|:|||: |:  :|  :|:|  | || |||
Db        601 KTKETKIIAKLGEGDDNVFVGSGTTEIDGGEGYDRVHYSR-GNYGALTIDATKETEQGSY 659

Qy        636 TVNRKVARGDIYHEVVKRQETKVGKRTETIQYRDYELRKVGYGYQSTDNLKSVEEVIGSQ 695
              |||| |   |  |||         || |  | |:|| :    :   || : | ||:|||:||:
Db        660 TVNRFVESGKALHEVTSTHTALVGNREEKIEYR-HSNNQHHAGYYTKDTLKAVEEIIGTS 718

Qy        696 FNDVFKGSKFNDIFHSGEGDDLLDGGAGDDRLFGGKGNDRLSGDEGDDLLDGGSGDDVLN 755
              ||:||||||||| |: |:| | :|| |:||||||||:| : |  ||| :||| |:|:|:
Db        719 HNDIFKGSKFNDAFNGGDGVDTIDGNDGNDRLFGGKGDDIIDGGNGDDFIDGGKGNDLLH 778

Qy        756 GGAGNDVYIFRKGDGNDTLYDGTGNDKLAFADANISDIMIERTKEGIIVKRNDHSGSINI 815
              || |:|::: |:|||||:: :  ||||:|:|:|: |:  |:    ::: |      : |
Db        779 GGKGDDIFVHRQGDGNDSITESEGNDKLSFSDSNLKDLTFEKVNHHLVI-TNTKQEKVTI 837

Qy        816 PRWY----ITSNLQNYQSNKTDHKIEQLIGKDGSYITSDQIDKILQDKKDGTVITSQELK 871
              |:      :|||  :: |  | |||::|||:|  ||| |||::::  |     |||
Db        838 QNWFREAEFAKTIQNYVATR-DDKIEEIIGQNGERITSKQVDELIE--KGNGKIAQSELT 894

Qy        872 KLADENKSQKLSASDIASSLNKLVGSMALFGTANSVSSNALQPITQPTQGILAPSV 927
              |: |  :  | |  | ::||:||: | :| ::| | | |      || :| ||:
Db        895 KVVDNYQLLKYS-RDASNSLDKLISSASAFTSSND-SRNVL---ASPT-SMLDPSL 944

RESULT 9
Q9EV28
ID   Q9EV28      PRELIMINARY;      PRT;   953 AA.
AC   Q9EV28;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Leukotoxin (LktA).
GN   LKTA.
OS   Pasteurella haemolytica.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC   Pasteurellaceae; Mannheimia.
OX   NCBI_TaxID=75985;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PH292, and PH296;
RX   MEDLINE=21101823; PubMed=11157953;
RA   Davies R.L., Whittam T.S., Selander R.K.;
RT   "Sequence Diversity and Molecular Evolution of the Leukotoxin (lktA)
RT   Gene in Bovine and Ovine Strains of Mannheimia (Pasteurella)
RT   haemolytica.";
RL   J. Bacteriol. 183:1394-1404(2001).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PH296;
RA   Davies R.L., Campbell S., Whittam T.S.;
RT   "Mosaic structure and molecular evolution of the leukotoxin operon
RT   (lktCABD) of Mannheimia (pasteurella) haemolytica, Mannheimia
RT   glucosida and Pasteurella trehalosi.";
RL   J. Bacteriol. 0:0-0(2001).
DR   EMBL; AF314515; AAG40299.1; -.
DR   EMBL; AF414141; AAL13281.1; -.
DR   InterPro; IPR001753; EnCoA_hydrtse.
DR   InterPro; IPR001343; Hemlysn_Ca_bind.
DR   InterPro; IPR003995; RtxA.
DR   Pfam; PF00353; hemolysinCabind; 5.
DR   Pfam; PF02382; RTX; 1.
DR   PRINTS; PR00313; CABNDNGRPT.
DR   PRINTS; PR01488; RTXTOXINA.
DR   PROSITE; PS00166; ENOYL_COA_HYDRATASE; 1.
SQ   SEQUENCE   953 AA;  102218 MW;  04AB1715B819E571 CRC64;

  Query Match             50.2%;  Score 2332.5;  DB 2;  Length 953;
  Best Local Similarity   49.6%;  Pred. No. 2.8e-96;
  Matches  474;  Conservative  180;  Mismatches  257;  Indels   45;  Gaps   18;

Qy          1 MSNINV-IKSN---IQAGLNST------------KSGLKNLYLAIPKD--YDPQKGGTLND 43
              ::||:  :||:    ::|||  |              |:| | : | |||| || || | |
Db          5 LTNISTNLKSSWLTAKSGLNRTGQSLAKAGQSLKTGAKKIILYIPKDYQYDTDKGNGLQD 64

Qy         44 FIKAADELGIARLAEEPNHTETAKKSVDTVNQFLSLTQTGIAISATKLEKFLQKHSTNKL 103
              :|||:||||      || |    |: |: |::  | ||: || :|| :|:| |||      |:
Db         65 LVKAAEELGIEVQKEESNDIAKAQTSLGTIHNVLGLTERGIVLSAPQLDKLLQK---TKV 121

Qy        104 AKGLDSVENIDRKLGKASNVLSTLSSFLGTALAGIELDSLIKKGDAAPDALAKASIDLIN 163
              : : | ||| :     |  |||: | ||: ||:||  : : ::      |||| ::| |
Db        122 GQAIGSTENITKGFSNAKTVLSGIQSILGSVLAGMDLDEAL-QNNSNELTLAKAGLELTN 180

Qy        164 EIIGNLSQSTQTIEAFSSQLAKLGSTISQAKGFSNIGNKLQNLN-FSKTNLGLEIITGLL 222
              :| |:: | :|::|| |: :|| :    || |::| ||: |: | ||:|||:|::|||
Db        181 SLIENIANSVKTLDAFGDQINQLGSKLQNVKGLSSLGEKLKGLSGFDKTSLGLDIVSGLL 240

Qy        223 SGISAGFALADKNASTGKKVAAGFELSNQVIGNVTKAISSYVLAQRVAAGLSTTGAVAAL 282
              || :|   ||||||| :|| ||||:|||:|||:|||:||||||||||:|| ||||
Db        241 SGATAALVLADKNASTSRKVGAGFELANQVVGNITKAVSSYILAQRVAAGLSSTGPVAAL 300

Qy        283 ITSSIMLAISPLAFMNAADKFNHANALDEFAKQFRKFGYDGDHLLAEYQRGVGTIEASLT 342
              | |:: ||||||:|  ||||||| :|: :|::|:| |||||:||||||||  |||:||:|
Db        301 IASTVSLAISPLSFAGIADKFNHAKSLESYAERFKKLGYDGDNLLAEYQRGTGTIDASVT 360

Qy        343 TISTALGAVSAGVSAAAVGSAVGTPIALLVAGVTGLISGILEASKQAMFESVANRLQGKI 402
              |:||| |:: ||||| || | :||||||:|:||:|| ||: |||||| |||::  ||
Db        361 AINTALAAIAGGVSAAAAGSVVASPIALLVSGITGVISTILQYSKQAMFEHVANKIHNKI 420

Qy        403 LEWEKQNGGQNYFDKGYDSRYAAYLANNLKFLSELNKELEAERVIAITQQRWDNNIGELA 462
              :|||| | |:|||: |||:|| | | :|:||| |||||:|||||||||||:|||||||:||
Db        421 VEWEKNNPGKNYFENGYDARYLANLQDNMKFLLNLNKELQAERVIAITQQQWDNNIGDLA 480

Qy        463 GITKLGERIKSGKAYADAFEDGKKVEAGSNITLDAKTGIIDISNSNGKKTQALHFTSPLL 522
              ||::|||:: ||||| ||||:|: ::|    : ||:  ||||:||:   ||| : | :|||
Db        481 GISRLGEKVLSGKAYVDAFEEGQHLKADKLVQLDSAKGIIDVSNTGEAKTQHILFRTPLL 540

Qy        523 TAGTESRERLTNGKYSYINKLKFGRVKNWQVTDGEASSKLDFSKVIQRV-------AETE 575
              | ||| |||:  ||| || ||   || :|::||| |||  | : |:||:
Db        541 TPGTEKRERVQTGKYEYITKLHINRVDSWKITDGAASSTFDLTNVVQRIGIELDNAGNVT 600

Qy        576 GTDEIGLIVNAKAGNDDIFVGQGKMNIDGGDGHDRVFYSKDGGFGNITVDGTSATEAGSY 635
              | |  :|   |:|::|||  ||||:|:|||: |:  :|  :|:|  | || |||
Db        601 KTKETKIIAKLGEGDDNVFVGSGTTEIDGGEGYDRVHYSR-GNYGALTIDATKETEQGSY 659

Qy        636 TVNRKVARGDIYHEVVKRQETKVGKRTETIQYRDYELRKVGYGYQSTDNLKSVEEVIGSQ 695
              |||| |   |  |||         || |  | |:|| :    :   || : | ||:|||:||:
Db        660 TVNRFVESGKALHEVTSTHTALVGNREEKIEYR-HSNNQHHAGYYTKDTLKAVEEIIGTS 718

Qy        696 FNDVFKGSKFNDIFHSGEGDDLLDGGAGDDRLFGGKGNDRLSGDEGDDLLDGGSGDDVLN 755
              ||:||||||||| |: |:| | :|| |:||||||||:| : |  ||| :||| |:|:|:
Db        719 HNDIFKGSKFNDAFNGGDGVDTIDGNDGNDRLFGGKGDDIIDGGNGDDFIDGGKGNDLLH 778

Qy        756 GGAGNDVYIFRKGDGNDTLYDGTGNDKLAFADANISDIMIERTKEGIIVKRNDHSGSINI 815
              || |:|::: |:|||||:: :  ||||:|:|:|: |:  |:    ::: |      : |
```

```
Db        779 GGKGDDIFVHRQGDGNDSITESEGNDKLSFSDSNLKDLTFEKVNHHLVI-TNTKQEKVTI 837

Qy        816 PRWY----ITSNLQNYQSNKTDHKIEQLIGKDGSYITSDQIDKILQDKKDGTVITSQELK 871
              |:        :||| : : | |||::||::| ||| |:|:::: |    |    ||
Db        838 QNWFREAEFAKTIQNYVATR-DDKIEEIIGQNGERITSKQVDELIE--KGNGKIAQSELT 894

Qy        872 KLADENKSQKLSASDIASSLNKLVGSMALFGTANSVSSNALQPITQPTQGILAPSV 927
              |: |  :  | |  | ::||:||: | : | ::| | | |     ||  :| ||:
Db        895 KVVDNYQLLKYS-RDASNSLDKLISSASAFTSSND-SRNVL---ASPT-SMLDPSL 944


RESULT 10
Q9ETG5
ID   Q9ETG5     PRELIMINARY;      PRT;   953 AA.
AC   Q9ETG5;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Leukotoxin.
GN   LKTA.
OS   Pasteurella haemolytica.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC   Pasteurellaceae; Mannheimia.
OX   NCBI_TaxID=75985;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PH284, PH388, and PH8;
RA   Davies R.L., Whittam T.S., Selander R.K.;
RT   "Sequence diversity and molecular evolution of the leukotoxin (lktA)
RT   gene in bovine and ovine strains of Mannheimia (Pasteurella)
RT   haemolytica.";
RL   J. Bacteriol. 0:0-0(2001).
DR   EMBL; AF314507; AAG40291.1; -.
DR   EMBL; AF314504; AAG40288.1; -.
DR   EMBL; AF314506; AAG40290.1; -.
DR   InterPro; IPR001753; EnCoA_hydrtse.
DR   InterPro; IPR001343; Hemlysn_Ca_bind.
DR   InterPro; IPR003995; RtxA.
DR   Pfam; PF00353; hemolysinCabind; 5.
DR   Pfam; PF02382; RTX; 1.
DR   PRINTS; PR00313; CABNDNGRPT.
DR   PRINTS; PR01488; RTXTOXINA.
DR   PROSITE; PS00166; ENOYL_COA_HYDRATASE; 1.
SQ   SEQUENCE   953 AA;  101997 MW;  D593D6A577C3ADE9 CRC64;

  Query Match            50.2%;  Score 2332;  DB 2;  Length 953;
  Best Local Similarity   50.1%;  Pred. No. 3e-96;
  Matches  465;  Conservative  175;  Mismatches  263;  Indels   26;  Gaps   13;

Qy          8 KSNIQAGLNSTKSGLKNLYLAIPKD--YDPQKGGTLNDFIKAADELGIARLAEEPNHTET 65
              :|  ||| :| |:| |  : | ||::  || ::| | | :|||:||||     || |:  |
Db         28 QSLTQAG-SSLKTGAKKIILYIPQNYQYDTEQGNGLQDLVKAAEELGIEVQREERNNIAT 86

Qy         66 AKKSVDTVNQFLSLTQTGIAISATKLEKFLQKHSTNKLAKGLDSVENIDRKLGKASNVLS 125
              |: |: |:    : ||: || :|| :::| |||      |   : | | |:|  :   || |||
Db         87 AQTSLGTIQTAIGLTERGIVLSAPQIDKLLQK---TKAGQALGSAESIVQNANKAKTVLS 143

Qy        126 TLSSFLGTALAGIELDSLIKKGDAAPDALAKASIDLINEIIGNLSQSTQTIEAFSSQLAK 185
               : | ||: |||::||   : : ::    ||||| ::| | :| |:: | :|:: |  |:::
Db        144 GIQSILGSVLAGMDLDEAL-QNNSNQHALAKAGLELTNSLIENIANSVKTLDEFGEQISQ 202

Qy        186 LGSTISQAKGFSNIGNKLQNL-NFSKTNLGLEIITGLLSGISAGFALADKNASTGKKVAA 244
               || :    ||     :|:||:|:    |  |||::|:||||| :|   |||||||| ||| |
Db        203 FGSKLQNIKGLGTLGDKLKNIGGLDKAGLGLDVISGLLSGATAALVLADKNASTAKKVGA 262

Qy        245 GFELSNQVIGNVTKAISSYVLAQRVAAGLSTTGAVAALITSSIMLAISPLAFMNAADKFN 304
              ||||:|||:||:|||:|||:||||||||||:|| ||||| |:: ||||||| |||||
Db        263 GFELANQVVGNITKAVSSYILAQRVAAGLSSTGPVAALIASTVSLAISPLAFAGIADKFN 322

Qy        305 HANALDEFAKQFRKFGYDGDHLLAEYQRGVGTIEASLTTISTALGAVSAGVSAAAVGSAV 364
              || :|: :|::|:| |||||:|||||||| |||:|:| |:|| |:: ||||| || :
Db        323 HAKSLESYAERFKKLGYDGDNLLAEYQRGTGTIDASVTAINTALAAIAGGVSAAAAGSVI 382

Qy        365 GTPIALLVAGVTGLISGILEASKQAMFESVANRLQGKILEWEKQNGGQNYFDKGYDSRYA 424
               :||||||:|:||:||  ||: |||||| |||::  ||:||||  | |:|||: |||:||
Db        383 ASPIALLVSGITGVISTILQYSKQAMFEHVANKIHNKIVEWEKNNHGKNYFENGYDARYL 442

Qy        425 AYLANNLKFLSELNKELEAERVIAITQQRWDNNIGELAGITKLGERIKSGKAYADAFEDG 484
              | | :|:|||  |||||:||||||||||:||||||:||||::|||:: ||||| ||||:|
Db        443 ANLQDNMKFLLNLNKELQAERVIAITQQQWDNNIGDLAGISRLGEKVLSGKAYVDAFEEG 502

Qy        485 KKVEAGSNITLDAKTGIIDISNSNGKKTQALHFTSPLLTAGTESRERLTNGKYSYINKLK 544
              | ::|    :  ||:   ||||:|||   |||  :  | :|||| |||  |||:  ||| || ||
Db        503 KHIKADKLVQLDSANGIIDVSNSGKAKTQHILFRTPLLTPGTEHRERVQTGKYEYITKLN 562

Qy        545 FGRVKNWQVTDGEASSKLDFSKVIQRV-------AETEGTDEIGLIVNAKAGNDDIFVGQ 597
                  || :|::||| |||  | : |:||:                 | |  :|     |:|::|||
Db        563 INRVDSWKITDGAASSTFDLTNVVQRIGIELDNAGNVTKTKETKIIAKLGEGDDNVFVGS 622

Qy        598 GKMNIDGGDGHDRVFYSKDGGFGNITVDGTSATEAGSYTVNRKVARGDIYHEVVKRQETK 657
              |     ||||:|:||| ||: | :| :|:| |  || |||||| |   |   |||
Db        623 GTTEIDGGEGYDRVHYSR-GNYGALTIDATKETEQGSYTVNRFVETGKALHEVTSTHTAL 681

Qy        658 VGKRTETIQYRDYELRKVGYGYQSTDNLKSVEEVIGSQFNDVFKGSKFNDIFHSGEGDDL 717
              || |  |  |:|| :    :    || :  | ||:|||:||:  ||:||||||| |: |:| |
Db        682 VGNREEKIEYR-HSNNQHHAGYYTKDTLKAVEEIIGTSHNDIFKGSKFNDAFNGGDGVDT 740

Qy        718 LDGGAGDDRLFGGKGNDRLSGDEGDDLLDGGSGDDVLNGGAGNDVYIFRKGDGNDTLYDG 777
              :||  |:|||||||:| | |     |||  |:|:|:|| |:|::: ||||||  : |
Db        741 IDGNDGNDRLFGGKGDDILDGGNGDDFIDGGKGNDLLHGGKGDDIFVHRKGDGNDIITDS 800

Qy        778 TGNDKLAFADANISDIMIERTKEGIIVKRNDHSGSINIPRWY----ITSNLQNYQSNKTD 833
              |||||:|:|:|: |:  |:  |   :::  |      :  |  |:           : ||:: | |
Db        801 DGNDKLSFSDSNLKDLTFEKVKHNLVI-TNSKKEKVTIQDWFREADFAKEVPNYKATK-D 858

Qy        834 HKIEQLIGKDGSYITSDQIDKILQDKKDGTVITSQELKKLADENKSQKLSASDIASSLNK 893
              |||::||::|  ||| |:| ::    |    ||  ||  |: |  :  | |  :: :||:|
Db        859 EKIEEIIGQNGERITSKQVDDLI--AKGNGKITQDELSKVVDNYELLKHS-KNVTNSLDK 915

Qy        894 LVGSMALFGTANSVSSNALQPITQPTQGI 922
              |: |:: | ::|  :  : | :  | :
Db        916 LISSVSAFTSSNDSRNVLVAPTSMLDQSL 944


RESULT 11
Q9EV31
ID   Q9EV31     PRELIMINARY;      PRT;   953 AA.
AC   Q9EV31;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Leukotoxin.
GN   LKTA.
OS   Pasteurella haemolytica.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC   Pasteurellaceae; Mannheimia.
OX   NCBI_TaxID=75985;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PH588;
RX   MEDLINE=21101823; PubMed=11157953;
RA   Davies R.L., Whittam T.S., Selander R.K.;
RT   "Sequence Diversity and Molecular Evolution of the Leukotoxin (lktA)
RT   Gene in Bovine and Ovine Strains of Mannheimia (Pasteurella)
RT   haemolytica.";
RL   J. Bacteriol. 183:1394-1404(2001).
DR   EMBL; AF314510; AAG40294.1; -.
DR   InterPro; IPR001753; EnCoA_hydrtse.
DR   InterPro; IPR001343; Hemlysn_Ca_bind.
DR   InterPro; IPR003995; RtxA.
DR   Pfam; PF00353; hemolysinCabind; 5.
DR   Pfam; PF02382; RTX; 1.
DR   PRINTS; PR00313; CABNDNGRPT.
DR   PRINTS; PR01488; RTXTOXINA.
DR   PROSITE; PS00166; ENOYL_COA_HYDRATASE; 1.
SQ   SEQUENCE   953 AA;  102160 MW;  A189BF80754A7907 CRC64;
```

```
Query Match            50.1%;  Score 2329.5;  DB 2;  Length 953;
Best Local Similarity  50.1%;  Pred. No. 3.9e-96;
Matches  469;  Conservative  178;  Mismatches  259;  Indels   31;  Gaps   16;

Qy          5 NVIKSNIQAGLNSTKSGLKNLYLAIPKD--YDPQKGGTLNDFIKAADELGIARLAEEPNH 62
Db         25 NAGQSLAKAG-QSLKTGAKKIILYIPKDYQYDTEKGNGLQDLVKAAQELGIEVQKEEGND 83

Qy         63 TETAKKSVDTVNQFLSLTQTGIAISATKLEKFLQKHSTNKLAKGLDSVENIDRKLGKASN 122
Db         84 IAKAQTSLGTIQNVLGLTERGIVLSAPQLDKLLQK---TKVGQAIGSAENLTKGFSNAKT 140

Qy        123 VLSTLSSFLGTALAGIELDSLIKKGDAAPDALAKASIDLINEIIGNLSQSTQTIEAFSSQ 182
Db        141 VLSGIQSILGSVLAGMDLDEALQK-NSNELTLAKAGLELTNSLIENIANSVKTLDAFGDQ 199

Qy        183 LAKLGSTISQAKGFSNIGNKLQNLN-FSKTNLGLEIITGLLSGISAGFALADKNASTGKK 241
Db        200 INQLGSKLQNVKGLSSLGDKLKGLSGFDKTSLGLDVVSGLLSGATAALVLADKNASTSRK 259

Qy        242 VAAGFELSNQVIGNVTKAISSYVLAQRVAAGLSTTGAVAALITSSIMLAISPLAFMNAAD 301
Db        260 VGAGFELANQVVGNITKAVSSYILAQRVAAGLSSTGPVAALIASTVSLAISPLAFAGIAD 319

Qy        302 KFNHANALDEFAKQFRKFGYDGDHLLAEYQRGVGTIEASLTTISTALGAVSAGVSAAAVG 361
Db        320 KFNHAKSLESYAERFKKLGYDGDNLLAEYQRGTGTIDASVTAINTALAAIAGGVSAAAAG 379

Qy        362 SAVGTPIALLVAGVTGLISGILEASKQAMFESVANRLQGKILEWEKQNGGQNYFDKGYDS 421
Db        380 SVIASPIALLVSGITGVISTILQYSKQAMFEHVANKIHNKIVEWEKNNGGKNYFENGYDA 439

Qy        422 RYAAYLANNLKFLSELNKELEAERVIAITQQRWDNNIGELAGITKLGERIKSGKAYADAF 481
Db        440 RYLANLQDNMKFLLNLNKELQAERVIAITQQQWDSNIGDLAGISRLGEKVLSGKAYVDAF 499

Qy        482 EDGKKVEAGSNITLDAKTGIIDISNSNGKKTQALHFTSPLLTAGTESRERLTNGKYSYIN 541
Db        500 EEGQHLKADKLVQLDSAKGIIDVSNTGEAKTQHILFRTPLLTPGTEKRERVQTGKYEYIT 559

Qy        542 KLKFGRVKNWQVTDGEASSKLDFSKVIQRVA-------ETEGTDEIGLIVNAKAGNDDIF 594
Db        560 KLHINRVDSWQIKDGAASSTFDLTNVVQRIGVELDHAENVIKTKETKIVATLGDGDDNVF 619

Qy        595 VGQGKMNIDGGDGHDRVFYSKDGGFGNITVDGTSATEAGSYTVNRKVARGDIYHEVVKRQ 654
Db        620 VGSGTTEIDGGEGYDRVHYSR-GNYGALTIDATKETEQGSYTVNRFVESGKALHEVTSTH 678

Qy        655 ETKVGKRTETIQYRDYELRKVGYGYQSTDNLKSVEEVIGSQFNDVFKGSKFNDIFHSGEG 714
Db        679 TALVGNREEKIEYR-HSNNQHHAGYYTKDTLKAVEEIIGTSHNDIFKGSKFNDAFNGGDG 737

Qy        715 DDLLDGGAGDDRLFGGKGNDRLSGDEGDDLLDGGSGDDVLNGGAGNDVYIFRKGDGNDTL 774
Db        738 VDTIDGNDGNDRLFGGKGDDIIDGGNGDDFIDGGKGNDLLHGGKGDDIFVHRQGDGNDSI 797

Qy        775 YDGTGNDKLAFADANISDIMIERTKEGIIVKRNDHSGSINIPRWY----ITSNLQNYQSN 830
Db        798 TESEGNDKLSFSDSNLKDLTFEKVNHHLVI-TNTKQEKVTIQNWFREAEFAKTIQNYVAT 856

Qy        831 KTDHKIEQLIGKDGSYITSDQIDKILQDKKDGTVITSQELKKLADENKSQKLSASDIASS 890
Db        857 R-DDKIEEIIGQNGERITSKQVDELIE--KGNGKIAQSELTKVVDNYQLLKYS-RDASNS 912

Qy        891 LNKLVGSMALFGTANSVSSNALQPITQPTQGILAPSV 927
Db        913 LDKLISSASAFTSSND-SRNVL---ASPT-SMLDPSL 944

RESULT 12
Q9EV27
ID   Q9EV27      PRELIMINARY;      PRT;   953 AA.
AC   Q9EV27;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Leukotoxin.
GN   LKTA.
OS   Pasteurella haemolytica.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC   Pasteurellaceae; Mannheimia.
OX   NCBI_TaxID=75985;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PH392;
RX   MEDLINE=21101823; PubMed=11157953;
RA   Davies R.L., Whittam T.S., Selander R.K.;
RT   "Sequence Diversity and Molecular Evolution of the Leukotoxin (lktA)
RT   Gene in Bovine and Ovine Strains of Mannheimia (Pasteurella)
RT   haemolytica.";
RL   J. Bacteriol. 183:1394-1404(2001).
DR   EMBL; AF314516; AAG40300.1; -.
DR   InterPro; IPR001753; EnCoA_hydrtse.
DR   InterPro; IPR001343; Hemlysn_Ca_bind.
DR   InterPro; IPR003995; RtxA.
DR   Pfam; PF00353; hemolysinCabind; 5.
DR   Pfam; PF02382; RTX; 1.
DR   PRINTS; PR00313; CABNDNGRPT.
DR   PRINTS; PR01488; RTXTOXINA.
DR   PROSITE; PS00166; ENOYL_COA_HYDRATASE; 1.
SQ   SEQUENCE   953 AA;  102230 MW;  2B686808EB370090 CRC64;

Query Match            50.1%;  Score 2329.5;  DB 2;  Length 953;
Best Local Similarity  49.6%;  Pred. No. 3.9e-96;
Matches  474;  Conservative  180;  Mismatches  257;  Indels   45;  Gaps   18;

Qy          1 MSNINV-IKSN---IQAGLNST-----------KSGLKNLYLAIPKD--YDPQKGGTLND 43
Db          5 LTNISTNLKSSWLTAKSGLNRTGQSLAKAGQSLKTGAKKIILYIPKDYQYDTDKGNGLQD 64

Qy         44 FIKAADELGIARLAEEPNHTETAKKSVDTVNQFLSLTQTGIAISATKLEKFLQKHSTNKL 103
Db         65 LVKAAEELGIEVQKEESNDIAKAQTSLGTIHNVLGLTERGIVLSAPQLDKLLQK---TKV 121

Qy        104 AKGLDSVENIDRKLGKASNVLSTLSSFLGTALAGIELDSLIKKGDAAPDALAKASIDLIN 163
Db        122 GQAIGSTENITKGFSNAKIVLSGIQSILGSVLAGMDLDEAL-QNNSNELTLAKAGLELTN 180

Qy        164 EIIGNLSQSTQTIEAFSSQLAKLGSTISQAKGFSNIGNKLQNLN-FSKTNLGLEIITGLL 222
Db        181 SLIENIANSVKTLDAFGDQINQLGSKLQNVKGLSSLGEKLKGLSGFDKTSLGLDIVSGLL 240

Qy        223 SGISAGFALADKNASTGKKVAAGFELSNQVIGNVTKAISSYVLAQRVAAGLSTTGAVAAL 282
Db        241 SGATAALVLADKNASTSRKVGAGFELANQVVGNITKAVSSYILAQRVAAGLSSTGPVAAL 300

Qy        283 ITSSIMLAISPLAFMNAADKFNHANALDEFAKQFRKFGYDGDHLLAEYQRGVGTIEASLT 342
Db        301 IASTVSLAISPLSFAGIADKFNHAKSLESYAERFKKLGYDGDNLLAEYQRGTGTIDASVT 360

Qy        343 TISTALGAVSAGVSAAAVGSAVGTPIALLVAGVTGLISGILEASKQAMFESVANRLQGKI 402
Db        361 AINTALAAIAGGVSAAAAGSVVASPIALLVSGITGVISTILQYSKQAMFEHVANKIHNKI 420

Qy        403 LEWEKQNGGQNYFDKGYDSRYAAYLANNLKFLSELNKELEAERVIAITQQRWDNNIGELA 462
Db        421 VEWEKNNPGKNYFENGYDARYLANLQDNMKFLLNLNKELQAERVIAITQQQWDNNIGDLA 480

Qy        463 GITKLGERIKSGKAYADAFEDGKKVEAGSNITLDAKTGIIDISNSNGKKTQALHFTSPLL 522
Db        481 GISRLGEKVLSGKAYVDAFEEGQHLKADKLVQLDSAKGIIDVSNTGEAKTQHILFRTPLL 540

Qy        523 TAGTESRERLTNGKYSYINKLKFGRVKNWQVTDGEASSKLDFSKVIQRV-------AETE 575
Db        541 TPGTEKRERVQTGKYEYITKLHINRVDSWKITDGAASSTFDLTNVVQRIGIELDNAGNVT 600
```

```
Qy        576 GTDEIGLIVNAKAGNDDIFVGQGKMNIDGGDGHDRVFYSKDGGFGNITVDGTSATEAGSY 635
Db        601 KTKETKIIAKLGEGDDNVFVGSGTTEIDGGEGYDRVHYSR-GNYGALTIDATKETEQGSY 659

Qy        636 TVNRKVARGDIYHEVVKRQETKVGKRTETIQYRDYELRKVGYGYQSTDNLKSVEEVIGSQ 695
Db        660 TVNRFVESGKALHEVTSTHTALVGNREEKIEYR-HSNNQHHAGYYTKDTLKAVEEIIGTS 718

Qy        696 FNDVFKGSKFNDIFHSGEGDDLLDGGAGDDRLFGGKGNDRLSGDEGDDLLDGGSGDDVLN 755
Db        719 HNDIFKGSKFNDAFNGGDGVDTIDGNDGNDRLFGGKGDDIIDGGNGDDFIDGGKGNDLLH 778

Qy        756 GGAGNDVYIFRKGDGNDTLYDGTGNDKLAFADANISDIMIERTKEGIIVKRNDHSGSINI 815
Db        779 GGKGDDIFVHRQGDGNDSITESEGNDKLSFSDSNLKDLTFEKVNHHLVI-TNTKQEKVTI 837

Qy        816 PRWY----ITSNLQNYQSNKTDHKIEQLIGKDGSYITSDQIDKILQDKKDGTVITSQELK 871
Db        838 QNWFREAEFAKTIQNYVATR-DDKIEEIIGQNGERITSKQVDELIE--KGNGKIAQSELT 894

Qy        872 KLADENKSQKLSASDIASSLNKLVGSMALFGTANSVSSNALQPITQPTQGILAPSV 927
Db        895 KVVDNYQLLKYS-RDASNSLDKLISSASAFTSSND-SRNVL---ASPT-SMLDPSL 944
```

RESULT 13
Q9EV29

```
ID   Q9EV29      PRELIMINARY;      PRT;   953 AA.
AC   Q9EV29;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Leukotoxin.
GN   LKTA.
OS   Pasteurella haemolytica.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC   Pasteurellaceae; Mannheimia.
OX   NCBI_TaxID=75985;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PH278;
RX   MEDLINE=21101823; PubMed=11157953;
RA   Davies R.L., Whittam T.S., Selander R.K.;
RT   "Sequence Diversity and Molecular Evolution of the Leukotoxin (lktA)
RT   Gene in Bovine and Ovine Strains of Mannheimia (Pasteurella)
RT   haemolytica.";
RL   J. Bacteriol. 183:1394-1404(2001).
DR   EMBL; AF314514; AAG40298.1; -.
DR   InterPro; IPR001753; EnCoA_hydrtse.
DR   InterPro; IPR001343; Hemlysn_Ca_bind.
DR   InterPro; IPR003995; RtxA.
DR   Pfam; PF00353; hemolysinCabind; 5.
DR   Pfam; PF02382; RTX; 1.
DR   PRINTS; PR00313; CABNDNGRPT.
DR   PRINTS; PR01488; RTXTOXINA.
DR   PROSITE; PS00166; ENOYL_COA_HYDRATASE; 1.
SQ   SEQUENCE   953 AA;  102147 MW;  11600FDA7849A1CA CRC64;

  Query Match             50.1%;  Score 2329;  DB 2;  Length 953;
  Best Local Similarity   49.9%;  Pred. No. 4.1e-96;
  Matches  465;  Conservative  175;  Mismatches  266;  Indels   26;  Gaps   13;

Qy          5 NVIKSNIQAGLNSTKSGLKNLYLAIPKD--YDPQKGGTLNDFIKAADELGIARLAEEPNH 62
Db         25 NAGQSLAKAG-QSLKTGAKKIILYIPQNYQYDTEQGNGLQDLVKAAEELGIEVQREERNN 83

Qy         63 TETAKKSVDTVNQFLSLTQTGIAISATKLEKFLQKHSTNKLAKGLDSVENIDRKLGKASN 122
Db         84 IATAQTSLGTIQTAIGLTERGIVLSAPQIDKLLQK---TKAGQALGSAESIVQNANKAKT 140

Qy        123 VLSTLSSFLGTALAGIELDSLIKKGDAAPDALAKASIDLINEIIGNLSQSTQTIEAFSSQ 182
Db        141 VLSGIQSILGSVLAGMDLDEAL-QNNSNQHALAKAGLELTNSLIENIANSVKTLDEFGEQ 199

Qy        183 LAKLGSTISQAKGFSNIGNKLQNL-NFSKTNLGLEIITGLLSGISAGFALADKNASTGKK 241
Db        200 ISQFGSKLQNIKGLGTLGDKLKNIGGLDKAGLGLDVISGLLSGATAALVLADKNASTAKK 259

Qy        242 VAAGFELSNQVIGNVTKAISSYVLAQRVAAGLSTTGAVAALITSSIMLAISPLAFMNAAD 301
Db        260 VGAGFELANQVVGNITKAVSSYILAQRVAAGLSSTGPVAALIASTVSLAISPLAFAGIAD 319

Qy        302 KFNHANALDEFAKQFRKFGYDGDHLLAEYQRGVGTIEASLTTISTALGAVSAGVSAAAVG 361
Db        320 KFNHAKSLESYAERFKKLGYDGDNLLAEYQRGTGTIDASVTAINTALAAIAGGVSAAAAG 379

Qy        362 SAVGTPIALLVAGVTGLISGILEASKQAMFESVANRLQGKILEWEKQNGGQNYFDKGYDS 421
Db        380 SVIASPIALLVSGITGVISTILQYSKQAMFEHVANKIHNKIVEWEKNNHGKNYFENGYDA 439

Qy        422 RYAAYLANNLKFLSELNKELEAERVIAITQQRWDNNIGELAGITKLGERIKSGKAYADAF 481
Db        440 RYLANLQDNMKFLLNLNKELQAERVIAITQQQWDNNIGDLAGISRLGEKVLSGKAYVDAF 499

Qy        482 EDGKKVEAGSNITLDAKTGIIDISNSNGKKTQALHFTSPLLTAGTESRERLTNGKYSYIN 541
Db        500 EEGKHIKADKLVQLDSANGIIDVSNSGKAKTQHILFRTPLLTPGTEHRERVQTGKYEYIT 559

Qy        542 KLKFGRVKNWQVTDGEASSKLDFSKVIQRV-------AETEGTDEIGLIVNAKAGNDDIF 594
Db        560 KLNINRVDSWKITDGAASSTFDLTNVVQRIGIELDNAGNVTKTKETKIIAKLGEGDDNVF 619

Qy        595 VGQGKMNIDGGDGHDRVFYSKDGGFGNITVDGTSATEAGSYTVNRKVARGDIYHEVVKRQ 654
Db        620 VGSGTTEIDGGEGYDRVHYSR-GNYGALTIDATKETEQGSYTVNRFVETGKALHEVTSTH 678

Qy        655 ETKVGKRTETIQYRDYELRKVGYGYQSTDNLKSVEEVIGSQFNDVFKGSKFNDIFHSGEG 714
Db        679 TALVGNREEKIEYR-HSNNQHHAGYYTKDTLKAVEEIIGTSHNDIFKGSKFNDAFNGGDG 737

Qy        715 DDLLDGGAGDDRLFGGKGNDRLSGDEGDDLLDGGSGDDVLNGGAGNDVYIFRKGDGNDTL 774
Db        738 VDTIDGNDGNDRLFGGKGDDILDGGNGDDFIDGGKGNDLLHGGKGDDIFVHRKGDGNDII 797

Qy        775 YDGTGNDKLAFADANISDIMIERTKEGIIVKRNDHSGSINIPRWY----ITSNLQNYQSN 830
Db        798 TDSDGNDKLSFSDSNLKDLTFEKVKHNLVI-TNSKKEKVTIQDWFREADFAKEVPNYKAT 856

Qy        831 KTDHKIEQLIGKDGSYITSDQIDKILQDKKDGTVITSQELKKLADENKSQKLSASDIASS 890
Db        857 K-DEKIEEIIGQNGERITSKQVDDLI--AKGNGKITQDELSKVVDNYELLKHS-KNVTNS 912

Qy        891 LNKLVGSMALFGTANSVSSNALQPITQPTQGI 922
Db        913 LDKLISSVSAFTSSNDSRNVLVAPTSMLDQSL 944
```

RESULT 14
Q9EV34

```
ID   Q9EV34      PRELIMINARY;      PRT;   953 AA.
AC   Q9EV34;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Leukotoxin.
GN   LKTA.
OS   Pasteurella haemolytica.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC   Pasteurellaceae; Mannheimia.
OX   NCBI_TaxID=75985;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PH56;
RX   MEDLINE=21101823; PubMed=11157953;
RA   Davies R.L., Whittam T.S., Selander R.K.;
```

```
RT   "Sequence Diversity and Molecular Evolution of the Leukotoxin (lktA)
RT   Gene in Bovine and Ovine Strains of Mannheimia (Pasteurella)
RT   haemolytica.";
RL   J. Bacteriol. 183:1394-1404(2001).
DR   EMBL; AF314505; AAG40289.1; -.
DR   InterPro; IPR001753; EnCoA_hydrtse.
DR   InterPro; IPR001343; Hemlysn_Ca_bind.
DR   InterPro; IPR003995; RtxA.
DR   Pfam; PF00353; hemolysinCabind; 5.
DR   Pfam; PF02382; RTX; 1.
DR   PRINTS; PR00313; CABNDNGRPT.
DR   PRINTS; PR01488; RTXTOXINA.
DR   PROSITE; PS00166; ENOYL_COA_HYDRATASE; 1.
SQ   SEQUENCE   953 AA;  102010 MW;  DE48B28EE0EB09FB CRC64;

  Query Match             50.1%;  Score 2326;  DB 2;  Length 953;
  Best Local Similarity   49.9%;  Pred. No. 5.6e-96;
  Matches  464;  Conservative  175;  Mismatches  264;  Indels   26;  Gaps   13;

Qy          8 KSNIQAGLNSTKSGLKNLYLAIPKD--YDPQKGGTLNDFIKAADELGIARLAEEPNHTET 65
Db         28 QSLTQAG-SSLKTGAKKIILYIPQNYQYDTEQGNGLQDLVKAAEELGIEVQREERNNIAT 86

Qy         66 AKKSVDTVNQFLSLTQTGIAISATKLEKFLQKHSTNKLAKGLDSVENIDRKLGKASNVLS 125
Db         87 AQTSLGTIQTAIGLTERGIVLSAPQIDKLLQK---TKAGQALGSAESIVQNANKAKTVLS 143

Qy        126 TLSSFLGTALAGIELDSLIKKGDAAPDALAKASIDLINEIIGNLSQSTQTIEAFSSQLAK 185
Db        144 GIQSILGSVLAGMDLDEAL-QNNSNQHALAKAGLELTNSLIENIANSVKTLDEFGEQISQ 202

Qy        186 LGSTISQAKGFSNIGNKLQNL-NFSKTNLGLEIITGLLSGISAGFALADKNASTGKKVAA 244
Db        203 FGSKLQNIKGLGTLGDKLKNIGGLDKAGLGLDVISGLLSGATAALVLADKNASTAKKVGA 262

Qy        245 GFELSNQVIGNVTKAISSYVLAQRVAAGLSTTGAVAALITSSIMLAISPLAFMNAADKFN 304
Db        263 GFELANQVVGNITKAVSSYILAQRVAAGLSSTGPVAALIASTVSLAISPLAFAGIADKFN 322

Qy        305 HANALDEFAKQFRKFGYDGDHLLAEYQRGVGTIEASLTTISTALGAVSAGVSAAAVGSAV 364
Db        323 HAKSLESYAERFKKLGYDGDNLLAEYQRGTGTIDASVTAINTALAAIAGGVSAAAAGSVI 382

Qy        365 GTPIALLVAGVTGLISGILEASKQAMFESVANRLQGKILEWEKQNGGQNYFDKGYDSRYA 424
Db        383 ASPIALLVSGITGVISTILQYSKQAMFEHVANKIHNKIVEWEKNNHGKNYFENGYDARYL 442

Qy        425 AYLANNLKFLSELNKELEAERVIAITQQRWDNNIGELAGITKLGERIKSGKAYADAFEDG 484
Db        443 ANLQDNMKFLLNLNKELQAERVIAITQQQWDNNIGDLAGISRLGEKVLSGKAYVDAFEEG 502

Qy        485 KKVEAGSNITLDAKTGIIDISNSNGKKTQALHFTSPLLTAGTESRERLTNGKYSYINKLK 544
Db        503 KHIKADKLVQLDSANGIIDVSNSGKAKTQHILFRTPLLTPGTEHRERVQTGKYEYITKLN 562

Qy        545 FGRVKNWQVTDGEASSKLDFSKVIQRV-------AETEGTDEIGLIVNAKAGNDDIFVGQ 597
Db        563 INRVDSWKITDGAASSTFDLTNVVQRIGIELDNAGNVTKTKETKIIAKLGEGDDNVFVGS 622

Qy        598 GKMNIDGGDGHDRVFYSKDGGFGNITVDGTSATEAGSYTVNRKVARGDIYHEVVKRQETK 657
Db        623 GTTEIDGGEGYDRVHYSR-GNYGALTIDATKEIEQGSYTVNRFVETGKALHEVTSTHTAL 681

Qy        658 VGKRTETIQYRDYELRKVGYGYQSTDNLKSVEEVIGSQFNDVFKGSKFNDIFHSGEGDDL 717
Db        682 VGNREEKIEYR-HSNNQHHAGYYTKDTLKAVEEIIGTSHNDIFKGSKFNDAFNGGDGVDT 740

Qy        718 LDGGAGDDRLFGGKGNDRLSGDEGDDLLDGGSGDDVLNGGAGNDVYIFRKGDGNDTLYDG 777
Db        741 IDGNDGNDRLFGGKGDDILDGGNGDDFIDGGKGNDLLHGGKGDDIFVHRKGDGNDIITDS 800

Qy        778 TGNDKLAFADANISDIMIERTKEGIIVKRNDHSGSINIPRWY----ITSNLQNYQSNKTD 833
Db        801 DGNDKLSFSDSNLKDLTFEKVKHNLVI-TNSKKEKVTIQDWFREADFAKEVPNYKATK-D 858

Qy        834 HKIEQLIGKDGSYITSDQIDKILQDKKDGTVITSQELKKLADENKSQKLSASDIASSLNK 893
Db        859 EKIEEIIGQNGERITSKQVDDLI--AKGNGKITQDELSKVVDNYELLKHS-KNVTNSLDK 915

Qy        894 LVGSMALFGTANSVSSNALQPITQPTQGI 922
Db        916 LISSVSAFTSSNDSRNVLVAPTSMLDQSL 944


RESULT 15
Q9EUE1
ID   Q9EUE1     PRELIMINARY;     PRT;   954 AA.
AC   Q9EUE1;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Leukotoxin.
GN   LKTA.
OS   Pasteurella haemolytica.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC   Pasteurellaceae; Mannheimia.
OX   NCBI_TaxID=75985;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PH202, and PH494;
RA   Davies R.L., Whittam T.S., Selander R.K.;
RT   "Sequence diversity and molecular evolution of the leukotoxin (lktA)
RT   gene in bovine and ovine strains of Mannheimia (Pasteurella)
RT   haemolytica.";
RL   J. Bacteriol. 0:0-0(2001).
DR   EMBL; AF314513; AAG40297.1; -.
DR   EMBL; AF314511; AAG40295.1; -.
DR   InterPro; IPR001753; EnCoA_hydrtse.
DR   InterPro; IPR001343; Hemlysn_Ca_bind.
DR   InterPro; IPR003995; RtxA.
DR   Pfam; PF00353; hemolysinCabind; 5.
DR   Pfam; PF02382; RTX; 1.
DR   PRINTS; PR00313; CABNDNGRPT.
DR   PRINTS; PR01488; RTXTOXINA.
DR   PROSITE; PS00166; ENOYL_COA_HYDRATASE; 1.
SQ   SEQUENCE   954 AA;  102410 MW;  02A60E52411711A2 CRC64;

  Query Match             49.9%;  Score 2320;  DB 2;  Length 954;
  Best Local Similarity   49.8%;  Pred. No. 1e-95;
  Matches  463;  Conservative  177;  Mismatches  248;  Indels   42;  Gaps   14;

Qy          5 NVIKSNIQAGLNSTKSGLKNLYLAIPKD--YDPQKGGTLNDFIKAADELGIARLAEEPNH 62
Db         25 NAGQSLAKAG-QSLKTGAKKIILYIPKDYQYDTDKGNGLQDLVKAAEELGIEVQKEESND 83

Qy         63 TETAKKSVDTVNQFLSLTQTGIAISATKLEKFLQKHSTNKLAKGLDSVENIDRKLGKASN 122
Db         84 IAKAQTSLGTIHNVLGLTERGIVLSAPQLDKLLQK---TKVGQAIGSTENITKGFSNAKT 140

Qy        123 VLSTLSSFLGTALAGIELDSLIKKGDAAPDALAKASIDLINEIIGNLSQSTQTIEAFSSQ 182
Db        141 VLSGIQSILGSVLAGMDLDEAL-QNNSNELTLAKAGLELTNSLIENIANSVKTLDAFGDQ 199

Qy        183 LAKLGSTISQAKGFSNIGNKLQNLN-FSKTNLGLEIITGLLSGISAGFALADKNASTGKK 241
Db        200 INQLGSKLQNVKGLSSLGEKLKGLSGFDKTSLGLDIVSGLLSGATAALVLADKNASTSRK 259

Qy        242 VAAGFELSNQVIGNVTKAISSYVLAQRVAAGLSTTGAVAALITSSIMLAISPLAFMNAAD 301
Db        260 VGAGFELANQVVGNITKAVSSYILAQRVAAGLSSTGPVAALIASTVSLAISPLSFAGIAD 319

Qy        302 KFNHANALDEFAKQFRKFGYDGDHLLAEYQRGVGTIEASLTTISTALGAVSAGVSAAAVG 361
Db        320 KFNHAKSLESYAERFKKLGYDGDNLLAEYQRGTGTIDASVTAINTALAAIAGGVSAAAAG 379

Qy        362 SAVGTPIALLVAGVTGLISGILEASKQAMFESVANRLQGKILEWEKQNGGQNYFDKGYDS 421
```

```
                | | :||||||:|:||:|| ||: |||||| |||::  ||:|||| | |:|||: |||:
Db        380 SVVASPIALLVSGITGVISTILQYSKQAMFEHVANKIHNKIVEWEKNNPGKNYFENGYDA 439

Qy        422 RYAAYLANNLKFLSELNKELEAERVIAITQQRWDNNIGELAGITKLGERIKSGKAYADAF 481
              || | | :|:|||  |||||:|||||||||:||||||:||||::|||:: ||||| |||
Db        440 RYLANLQDNMKFLLNLNKELQAERVIAITQQQWDNNIGDLAGISRLGEKVLSGKAYVDAF 499

Qy        482 EDGKKVEAGSNITLDAKTGIIDISNSNGKKTQALHFTSPLLTAGTESRERLTNGKYSYIN 541
              |:|| ::|     : ||:   ||||:|||   |||  :  | :||||  ||| |||:  ||| ||
Db        500 EEGKHLKADKLVQLDSANGIIDVSNSGKAKTQHILFRTPLLTPGTEKRERVQTGKYEYIT 559

Qy        542 KLKFGRVKNWQVTDGEASSKLDFSKVIQRV-------AETEGTDEIGLIVNAKAGNDDIF 594
              ||     || :|::||||||   | : |:||:             | |   :|      |:|::|
Db        560 KLNINRVDSWKITDGEASSTFDLTNVVQRIGIELDNAGNVTKTKETKIIAKLGEGDDNVF 619

Qy        595 VGQGKMNIDGGDGHDRVFYSKDGGFGNITVDGTSATEAGSYTVNRKVARGDIYHEVVKRQ 654
              || |     ||||:|:|||  ||:  |  :  :|:|  |    ||  |||||| |   |    |||
Db        620 VGSGTTEIDGGEGYDRVHYSR-GNYGALTIDATKETEQGSYTVNRFVETGKALHEVTSTH 678

Qy        655 ETKVGKRTETIQYRDYELRKVGYGYQSTDNLKSVEEVIGSQFNDVFKGSKFNDIFHSGEG 714
                  || | | |:||     |:   || : | | |:||:||:  ||:||||:||| |: |:|
Db        679 TALVGSREEKIEYRHSNNRQHA-GYYTKDTLTSIEEIIGTSHNDIFKGSQFNDAFNGGDG 737

Qy        715 DDLLDGGAGDDRLFGGKGNDRLSGDEGDDLLDGGSGDDVLNGGAGNDVYIFRKGDGNDTL 774
               | :||   |:||||||||:|  :  | :|||  :|||  |:|:|:||  |:|:::  |:||||||::
Db        738 VDTIDGNGGNDRLFGGKGDDIIDGGDGDDFIDGGKGNDLLHGGRGDDIFVHRQGDGNDSI 797

Qy        775 YDGTGNDKLAFADANISDIMIERTKEGIIVKRNDHSGSINIPRWY----ITSNLQNYQSN 830
               | :|:|:|:|:|:|: |:  |:   :::     |   |   |  |:      ::|| :
Db        798 TEAGGHDRLSFSDSNLKDLTFEKVNHHLVI-TNTKQEKVTIQNWFREEEFAKTVKNYVAT 856

Qy        831 KTDHKIEQLIGKDGSYITSDQIDKILQDKKD--------GTVITSQELKKLADENKSQKL 882
              : | |||::||::|  ||| |:|:::    ||                |: | ||| |
Db        857 R-DEKIEEIIGQNGERITSKQVDELIAKGKDNKIDKNDLANVVNSYELLK---------- 905

Qy        883 SASDIASSLNKLVGSMALFGTANSVSSNAL 912
              :: :: :||:||: |:: | ::| | | |
Db        906 NSRNVTNSLDKLISSVSSFTSSND-SRNVL 934
```

RESULT 16
Q9EV30
ID   Q9EV30   PRELIMINARY;   PRT;   953 AA.
AC   Q9EV30;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Leukotoxin.
GN   LKTA.
OS   Pasteurella haemolytica.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC   Pasteurellaceae; Mannheimia.
OX   NCBI_TaxID=75985;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PH196;
RX   MEDLINE=21101823; PubMed=11157953;
RA   Davies R.L., Whittam T.S., Selander R.K.;
RT   "Sequence Diversity and Molecular Evolution of the Leukotoxin (lktA)
RT   Gene in Bovine and Ovine Strains of Mannheimia (Pasteurella)
RT   haemolytica.";
RL   J. Bacteriol. 183:1394-1404(2001).
DR   EMBL; AF314512; AAG40296.1; -.
DR   InterPro; IPR001753; EnCoA_hydrtse.
DR   InterPro; IPR001343; Hemlysn_Ca_bind.
DR   InterPro; IPR003995; RtxA.
DR   Pfam; PF00353; hemolysinCabind; 5.
DR   Pfam; PF02382; RTX; 1.
DR   PRINTS; PR00313; CABNDNGRPT.
DR   PRINTS; PR01488; RTXTOXINA.
DR   PROSITE; PS00166; ENOYL_COA_HYDRATASE; 1.
SQ   SEQUENCE   953 AA; 102043 MW; 4E8F11490479A69A CRC64;

```
  Query Match             49.6%;  Score 2307;  DB 2;  Length 953;
  Best Local Similarity   49.3%;  Pred. No. 3.9e-95;
  Matches  458;  Conservative  179;  Mismatches  266;  Indels   26;  Gaps   13;

Qy          8 KSNIQAGLNSTKSGLKNLYLAIPKD--YDPQKGGTLNDFIKAADELGIARLAEEPNHTET 65
              :|  |||  :|  |:|  | : | ||::  || ::|  | | :|||:||||     || |  |
Db         28 QSLTQAG-SSLKTGAKKIILYIPQNYQYDTEQGNGLQDLVKAAEELGIEVQREERNDIAT 86

Qy         66 AKKSVDTVNQFLSLTQTGIAISATKLEKFLQKHSTNKLAKGLDSVENIDRKLGKASNVLS 125
              |: |: |:    : ||: || :|| :::| |||      |  : | | |:| :    || |||
Db         87 AQTSLGTIQTAIGLTERGIVLSAPQIDKLLQK---TKAGQALGSAESIVQNANKAKTVLS 143

Qy        126 TLSSFLGTALAGIELDSLIKKGDAAPDALAKASIDLINEIIGNLSQSTQTIEAFSSQLAK 185
               : | ||: |||::||  : : ::   |||| ::| | :| |:: | :|:: |  |:::
Db        144 GIQSILGSVLAGMDLDEAL-QNNSNQHALAKAGLELTNSLIENIANSVKTLDEFGEQISQ 202

Qy        186 LGSTISQAKGFSNIGNKLQNL-NFSKTNLGLEIITGLLSGISAGFALADKNASTGKKVAA 244
                || :    ||    :|:||:|:    |  |||::|:||||| :|   |||||||| ||| |
Db        203 FGSKLQNIKGLGTLGDKLKNIGGLDKAGLGLDVISGLLSGATAALVLADKNASTAKKVGA 262

Qy        245 GFELSNQVIGNVTKAISSYVLAQRVAAGLSTTGAVAALITSSIMLAISPLAFMNAADKFN 304
              ||||:||:|:||:|:|:||:|||||||||:||:|:|  ||||||||||  |::  |||||||  |||||
Db        263 GFELANQVVGNITKAVSSYILAQRVAAGLSSTGPVAALIASTVSLAISPLAFAGIADKFN 322

Qy        305 HANALDEFAKQFRKFGYDGDHLLAEYQRGVGTIEASLTTISTALGAVSAGVSAAAVGSAV 364
              || :|: :|::|:|  ||||:|||||||||||| |||:||:| |:||| |:: ||||| || :
Db        323 HAKSLESYAERFKKLGYDGDNLLAEYQRGTGTIDASVTAINTALAAIAGGVSAAAAGSVI 382

Qy        365 GTPIALLVAGVTGLISGILEASKQAMFESVANRLQGKILEWEKQNGGQNYFDKGYDSRYA 424
               :|||||:|:|:|:|| ||: |||||| |||::  ||:|||| | |:|||: |||:||
Db        383 ASPIALLVSGITGVISTILQYSKQAMFEHVANKIHNKIVEWEKNNQGKNYFENGYDARYL 442

Qy        425 AYLANNLKFLSELNKELEAERVIAITQQRWDNNIGELAGITKLGERIKSGKAYADAFEDG 484
              | | :|:|||  |||||:|||||||||:|:|:|||:||||::|||:: ||||| ||||:|
Db        443 ANLQDNMKFLLNLNKELQAERVIAITQQQWDSNIGDLAGISRLGEKVISGKAYVDAFEEG 502

Qy        485 KKVEAGSNITLDAKTGIIDISNSNGKKTQALHFTSPLLTAGTESRERLTNGKYSYINKLK 544
              | ::|     :  ||:   ||||:||:   |||: :| |:||| ||| |||:  |||| |||
Db        503 KHIKADKLVQLDSAKGIIDVSNTGEAKTQHILFRTPLLTPGTEKRERVQTGKYEYITKLN 562

Qy        545 FGRVKNWQVTDGEASSKLDFSKVIQRVA-------ETEGTDEIGLIVNAKAGNDDIFVGQ 597
                || :||: || ||| |  ||| | |:||:             |  |     ::     ||:|::|||
Db        563 INRVDSWQIKDGAASSTFDLTNVVQRIGIELDHAENVTKTKETKIVAKLGAGDDNVFVGS 622

Qy        598 GKMNIDGGDGHDRVFYSKDGGFGNITVDGTSATEAGSYTVNRKVARGDIYHEVVKRQETK 657
              |    ||||:|:||| ||: | :| :|:| |  |  |||||| |   |    |||
Db        623 GTTEIDGGEGYDRVHYSR-GNYGALTIDATKETVQGSYTVNRFVETGKALHEVTSTHTAL 681

Qy        658 VGKRTETIQYRDYELRKVGYGYQSTDNLKSVEEVIGSQFNDVFKGSKFNDIFHSGEGDDL 717
              || | | |:||     |:   || : | | ::||:|||:  ||:|:||||||| |  |:| |
Db        682 VGSREEKIEYRHSNNRQHA-GYYTKDTLTTIEEIIGTSHNDIFRGSKFNDAFQGGDGVDT 740

Qy        718 LDGGAGDDRLFGGKGNDRLSGDEGDDLLDGGSGDDVLNGGAGNDVYIFRKGDGNDTLYDG 777
              :||  |:|||||||:| : |   ||| :|||  |:|:|:||  |:|::: ||||||| :|
Db        741 IDGNDGNDRLFGGKGDDIIDGGNGDDFIDGGKGNDLLHGGRGDDIFVHRKGDGNDIISDS 800

Qy        778 TGNDKLAFADANISDIMIERTKEGIIVKRNDHSGSINIPRWY----ITSNLQNYQSNKTD 833
              |||||:|:|:|: |:  |:   :::   |  :   |  |:      :    :||:  |
Db        801 DGNDKLSFSDSNLKDLTFEKVKYNLVI-TNSNKEKVTIQNWFREADLAKEVHNYKAT-AD 858

Qy        834 HKIEQLIGKDGSYITSDQIDKILQDKKDGTVITSQELKKLADENKSQKLSASDIASSLNK 893
              |||::||::|  ||| |:| :::   |      ||   ||  |  |  :  | |   ::|:||:|
Db        859 EKIEEIIGQNGERITSKQVDDLIE--KGNGKITQDELSKAVDNYELLKHS-KNVTNSLDK 915

Qy        894 LVGSMALFGTANSVSSNALQPITQPTQGI 922
              |: |:: | ::|  : | :  | :
Db        916 LISSVSSFTSSNDSRNVLVAPASMLDQSL 944
```

RESULT 17
Q9EV22
ID   Q9EV22   PRELIMINARY;   PRT;   955 AA.
AC   Q9EV22;

```
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Leukotoxin.
GN   LKTA.
OS   Pasteurella trehalosi.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC   Pasteurellaceae; Pasteurella.
OX   NCBI_TaxID=47735;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PH68;
RX   MEDLINE=21101823; PubMed=11157953;
RA   Davies R.L., Whittam T.S., Selander R.K.;
RT   "Sequence Diversity and Molecular Evolution of the Leukotoxin (lktA)
RT   Gene in Bovine and Ovine Strains of Mannheimia (Pasteurella)
RT   haemolytica.";
RL   J. Bacteriol. 183:1394-1404(2001).
DR   EMBL; AF314526; AAG40310.1; -.
DR   InterPro; IPR001753; EnCoA_hydrtse.
DR   InterPro; IPR001343; Hemlysn_Ca_bind.
DR   InterPro; IPR003995; RtxA.
DR   Pfam; PF00353; hemolysinCabind; 5.
DR   Pfam; PF02382; RTX; 1.
DR   PRINTS; PR00313; CABNDNGRPT.
DR   PRINTS; PR01488; RTXTOXINA.
DR   PROSITE; PS00166; ENOYL_COA_HYDRATASE; 1.
SQ   SEQUENCE   955 AA;  102069 MW;  FDCF2E99B8DC394F CRC64;

  Query Match             49.2%;  Score 2286;  DB 2;  Length 955;
  Best Local Similarity   49.0%;  Pred. No. 3.4e-94;
  Matches  466;  Conservative  177;  Mismatches  264;  Indels   44;  Gaps   17;

Qy          5 NVIKSNIQA---GLN-----------STKSGLKNLYLAIPKD--YDPQKGGTLNDFIKAA 48
                | |:| : |    |||           : |:| | : | ||||  ||    |  | | :|||
Db         12 NGIRSTLTATRGGLNRAGQSLTQAGQTLKNGAKKIILYIPKDYKYDSGSGNGLQDLVKAA 71

Qy         49 DELGIARLAEEPNHTETAKKSVDTVNQFLSLTQTGIAISATKLEKFLQKHSTNKLAKGLD 108
              :||||      || |      |: |: |:    | ||: || :|| :|:| |||   ||: : |
Db         72 EELGIEVQKEEGNDIAKAQTSLGTIQNVLGLTERGIVLSAPQLDKLLQK---NKVGQALG 128

Qy        109 SVENIDRKLGKASNVLSTLSSFLGTALAGIELDSLIKKGDAAPDALAKASIDLINEIIGN 168
              | |:| :    :|  ||| : | ||: |||::||   : : ::      |||| ::| | :| |
Db        129 SSESIAQNFSQAKTVLSGVQSILGSVLAGMDLDEAL-QNESDQLTLAKAGLELTNSLIEN 187

Qy        169 LSQSTQTIEAFSSQLAKLGSTISQAKGFSNIGNKLQNL-NFSKTNLGLEIITGLLSGISA 227
              :: | ||::||| |::: || :    ||    :|:||:|:     |  |||::|:||||| :|
Db        188 IANSVQTLDAFSEQISQFGSKLQNVKGLGALGDKLKNIGGLDKAGLGLDVISGLLSGATA 247

Qy        228 GFALADKNASTGKKVAAGFELSNQVIGNVTKAISSYVLAQRVAAGLSTTGAVAALITSSI 287
                   ||||:|||  |||  |||||:|||:||:|||:|||:|||||||||:||  |||||  |::
Db        248 ALVLADKDASTAKKVGAGFELANQVVGNITKAVSSYILAQRVAAGLSSTGPVAALIASTV 307

Qy        288 MLAISPLAFMNAADKFNHANALDEFAKQFRKFGYDGDHLLAEYQRGVGTIEASLTTISTA 347
                   ||||||:|    ||||:  | :|: :|::|:| ||:|| ||||| | |||:||:| |:||
Db        308 ALAISPLSFAGIADKFDRAKSLENYAERFKKLGYEGDSLLAEYQHGTGTIDASVTAINTA 367

Qy        348 LGAVSAGVSAAAVGSAVGTPIALLVAGVTGLISGILEASKQAMFESVANRLQGKILEWEK 407
              | |:: ||||| || |  :||||||:|:||:|| ||: |||||| |||::   ||:||||
Db        368 LAAIAGGVSAAAAGSVVASPIALLVSGITGVISTILQYSKQAMFEHVANKIHNKIVEWEK 427

Qy        408 QNGGQNYFDKGYDSRYAAYLANNLKFLSELNKELEAERVIAITQQRWDNNIGELAGITKL 467
                   |||:|||: |||:|| | |  :|:|||  |||||:||||||||||:||:|||:||||:|
Db        428 NNGGKNYFENGYDARYLANLQDNMKFLLNLNKELQAERVIAITQQQWDSNIGDLAGISRL 487

Qy        468 GERIKSGKAYADAFEDGKKVEAGSNITLDAKTGIIDISNSNGKKTQALHFTSPLLTAGTE 527
              ||:: |||||  ||||:|: ::|     : ||:  ||||:||:    |||  :  | :||||  |||
Db        488 GEKVLSGKAYVDAFEEGQHLKADKLVQLDSAKGIIDVSNTGEAKTQHILFRTPLLTPGTE 547

Qy        528 SRERLTNGKYSYINKLKFGRVKNWQVTDGEASSKLDFSKVIQRVA-------ETEGTDEI 580
                 |||:  ||| |||  ||    ||  ||:|:  || |||   |  : |:||:
Db        548 KRERVQTGKYEYITKLHINRVDSWQIKDGAASSTFDLTNVVQRIGVELDHAENVIKTKET 607

Qy        581 GLIVNAKAGNDDIFVGQGKMNIDGGDGHDRVFYSKDGGFGNITVDGTSATEAGSYTVNRK 640
                ::      |:|::||| |     ||||:|:||| ||: | :| :|:| |    || ||||||
Db        608 KIVATLGDGDDNVFVGSGTTEIDGGEGYDRVHYSR-GNYGALTIDATKETEQGSYTVNRF 666

Qy        641 VARGDIYHEVVKRQETKVGKRTETIQYRDYELRKVGYGYQSTDNLKSVEEVIGSQFNDVF 700
               |   |  |||          || | |  |:|| :   :     || : | ||:|||:||: ||:|
Db        667 VESGKALHEVTSTHTALVGNREEKIEYR-HSNNQHHAGYYTKDTLKAVEEIIGTSHNDIF 725

Qy        701 KGSKFNDIFHSGEGDDLLDGGAGDDRLFGGKGNDRLSGDEGDDLLDGGSGDDVLNGGAGN 760
              |||||| |: |:| | :||  |:||||||||:| : |  |||  :|||  ||:|:|:|| |:
Db        726 KGSKFNDAFNGGDGVDTIDGNDGNDRLFGGKGDDIIDGGNGDDFIDGGKGNDLLHGGKGD 785

Qy        761 DVYIFRKGDGNDTLYDGTGNDKLAFADANISDIMIERTKEGIIVKRNDHSGSINIPRWY- 819
              |::: |:||||:: :  ||||:|:|:|: |:  |:    :::  |     : |  |:
Db        786 DIFVHRQGDGNDSITESEGNDKLSFSDSNLKDLTFEKVNHHLVI-TNTKQEKVTIQNWFR 844

Qy        820 ---ITSNLQNYQSNKTDHKIEQLIGKDGSYITSDQIDKILQDKKDGTVITSQELKKLADE 876
                   :||| : : | |||::||::|  |||  |:|::::  |     |      || |: |
Db        845 EAEFAKTIQNYVATR-DDKIEEIIGQNGERITSKQVDELIE--KGNGKIAQSELTKVVDN 901

Qy        877 NKSQKLSASDIASSLNKLVGSMALFGTANSVSSNALQPITQPTQGILAPSV 927
                 :  | |   | ::||:|| : |  : | :: :|  |  | |     ||  :| ||:
Db        902 YQLLKYS-RDASNSLDKLISSASAFTSSND-SRNVL---ASPT-SMLDPSL 946


RESULT 18
Q9EUD4
ID   Q9EUD4     PRELIMINARY;     PRT;   955 AA.
AC   Q9EUD4;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Leukotoxin.
GN   LKTA.
OS   Pasteurella trehalosi.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC   Pasteurellaceae; Pasteurella.
OX   NCBI_TaxID=47735;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PH254, PH246, and PH252;
RA   Davies R.L., Whittam T.S., Selander R.K.;
RT   "Sequence diversity and molecular evolution of the leukotoxin (lktA)
RT   gene in bovine and ovine strains of Mannheimia (Pasteurella)
RT   haemolytica.";
RL   J. Bacteriol. 0:0-0(2001).
DR   EMBL; AF314525; AAG40309.1; -.
DR   EMBL; AF314523; AAG40307.1; -.
DR   EMBL; AF314524; AAG40308.1; -.
DR   InterPro; IPR001753; EnCoA_hydrtse.
DR   InterPro; IPR001343; Hemlysn_Ca_bind.
DR   InterPro; IPR003995; RtxA.
DR   Pfam; PF00353; hemolysinCabind; 5.
DR   Pfam; PF02382; RTX; 1.
DR   PRINTS; PR00313; CABNDNGRPT.
DR   PRINTS; PR01488; RTXTOXINA.
DR   PROSITE; PS00166; ENOYL_COA_HYDRATASE; 1.
SQ   SEQUENCE   955 AA;  102127 MW;  074B2E4ADFFF57AA CRC64;

  Query Match             48.9%;  Score 2274;  DB 2;  Length 955;
  Best Local Similarity   48.7%;  Pred. No. 1.2e-93;
  Matches  463;  Conservative  180;  Mismatches  264;  Indels   44;  Gaps   17;

Qy          5 NVIKSNIQA---GLN-----------STKSGLKNLYLAIPKD--YDPQKGGTLNDFIKAA 48
                | |:| : |    |||           : |:| | : | ||||  ||    |  | | :|||
Db         12 NGIRSTLTATRGGLNRAGQSLTQAGQTLKNGAKKIILYIPKDYKYDSGSGNGLQDLVKAA 71

Qy         49 DELGIARLAEEPNHTETAKKSVDTVNQFLSLTQTGIAISATKLEKFLQKHSTNKLAKGLD 108
              :||||      || |      |: |: |:    | ||: || :|| :|:| |||   ||: : |
Db         72 EELGIEVQKEEGNDIAKAQTSLGTIQNVLGLTERGIVLSAPQLDKLLQK---NKVGQALG 128

Qy        109 SVENIDRKLGKASNVLSTLSSFLGTALAGIELDSLIKKGDAAPDALAKASIDLINEIIGN 168
              | |:| :    :|  ||| : | ||: |||::||   : : ::      |||| ::| | :| |
```

```
Db        129 SSESIAQNFSQAKTVLSGVQSILGSVLAGMDLDEAL-QNESDQLTLAKAGLELTNSLIEN 187

Qy        169 LSQSTQTIEAFSSQLAKLGSTISQAKGFSNIGNKLQNL-NFSKTNLGLEIITGLLSGISA 227
Db        188 IANSVQTLDAFSEQISQFGSKLQNVKGLGALGDKLKNIGGLDKAGLGLDVISGLLSGATA 247

Qy        228 GFALADKNASTGKKVAAGFELSNQVIGNVTKAISSYVLAQRVAAGLSTTGAVAALITSSI 287
Db        248 ALVLADKDASTAKKVGAGFELANQVVGNITKAVSSYILAQRVAAGLSSTGPVAALIASTV 307

Qy        288 MLAISPLAFMNAADKFNHANALDEFAKQFRKFGYDGDHLLAEYQRGVGTIEASLTTISTA 347
Db        308 ALAISPLSFAGIADKFDRAKSLENYAERFKKLGYEGDSLLAEYQHGTGTIDASVTAINTA 367

Qy        348 LGAVSAGVSAAAVGSAVGTPIALLVAGVTGLISGILEASKQAMFESVANRLQGKILEWEK 407
Db        368 LAAIAGGVSAAAAGSVVASPIALLVSGITGVISTILQYSKQAMFEHVANKIHNKIVEWEK 427

Qy        408 QNGGQNYFDKGYDSRYAAYLANNLKFLSELNKELEAERVIAITQQRWDNNIGELAGITKL 467
Db        428 NNGGKNYFENGYDARYLANLQDNMKFLLNLNKELQAERVIAITQQQWDSNIGDLAGISRL 487

Qy        468 GERIKSGKAYADAFEDGKKVEAGSNITLDAKTGIIDISNSNGKKTQALHFTSPLLTAGTE 527
Db        488 GEKVLSGKAYVDAFEEGQHLKADKLVQLDSAKGIIDVSNTGEAKTQHILFRTPLLTPGTE 547

Qy        528 SRERLTNGKYSYINKLKFGRVKNWQVTDGEASSKLDFSKVIQRVA-------ETEGTDEI 580
Db        548 KRERVQTGKYEYITKLHINRVDSWQIKDGAASSTFDLTNVVQRIGVELDHAENVIKTKET 607

Qy        581 GLIVNAKAGNDDIFVGQGKMNIDGGDGHDRVFYSKDGGFGNITVDGTSATEAGSYTVNRK 640
Db        608 KIVATLGDGDDNVFVGSGTTEIDGGEGYDRVHYSR-GNYGALTIDATKETEQGSYTVNRF 666

Qy        641 VARGDIYHEVVKRQETKVGKRTETIQYRDYELRKVGYGYQSTDNLKSVEEVIGSQFNDVF 700
Db        667 VESGKALHEVTSTHTALVGNREEKIEYR-HSNNQHHAGYYTKDTLKAVEEIIGTSHNDIF 725

Qy        701 KGSKFNDIFHSGEGDDLLDGGAGDDRLFGGKGNDRLSGDEGDDLLDGGSGDDVLNGGAGN 760
Db        726 KGSKFNDAFNGGDGVDTIDGNDGNDRLFGGKGDDIIDGGNGDDFIDGGKGNDLLHGGKGD 785

Qy        761 DVYIFRKGDGNDTLYDGTGNDKLAFADANISDIMIERTKEGIIVKRNDHSGSINIPRWY- 819
Db        786 DIFVHRQGDGNDSITESEGNDKLSFSDSNLKDLTFEKVNHHLVI-TNTKQEKVTIQNWFR 844

Qy        820 ---ITSNLQNYQSNKTDHKIEQLIGKDGSYITSDQIDKILQDKKDGTVITSQELKKLADE 876
Db        845 EAEFAKTIRNYVATR-DDKIEEIIGQNGERITSKQVDELIEKGKG--KIDKSDLSQVVDN 901

Qy        877 NKSQKLSASDIASSLNKLVGSMALFGTANSVSSNALQPITQPTQGILAPSV 927
Db        902 YQLLKYS-RDASNSLDKLISSASAFTSSND-SRNVL---ASPT-SMLDPSL 946
```

RESULT 19
Q93NP1

```
ID   Q93NP1      PRELIMINARY;      PRT;    956 AA.
AC   Q93NP1;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   RTX toxin IIA.
OS   Actinobacillus pleuropneumoniae (Haemophilus pleuropneumoniae).
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC   Pasteurellaceae; Actinobacillus.
OX   NCBI_TaxID=715;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Shin S.J., Park J.Y., Choi I.S., Shin N.R., Yoo H.S.;
RT   "Cloning and Sequencing of Apx IIA from Actinobacillus
RT   pleuropneumoniae.";
RL   Submitted (MAR-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF363362; AAK50052.1; -.
DR   InterPro; IPR001343; Hemlysn_Ca_bind.
DR   InterPro; IPR003995; RtxA.
DR   Pfam; PF00353; hemolysinCabind; 5.
DR   Pfam; PF02382; RTX; 1.
DR   PRINTS; PR00313; CABNDNGRPT.
DR   PRINTS; PR01488; RTXTOXINA.
DR   PROSITE; PS00330; HEMOLYSIN_CALCIUM; 1.
SQ   SEQUENCE   956 AA;  102495 MW;  9F8EB14EC76628D5 CRC64;

  Query Match             48.0%;  Score 2229.5;  DB 2;  Length 956;
  Best Local Similarity   47.7%;  Pred. No. 1.1e-91;
  Matches  453;  Conservative  187;  Mismatches  256;  Indels   53;  Gaps   15;

Qy          1 MSNINV--IKSNIQAGL---------------------NSTKSGLKNLYLAIPKDYDPQ 36
Db          1 MSKITLSSLKSSLQQGLKNGKNKLNQAGTTLKNGLTQTGHSLQNGAKKLILYIPQGYDSG 60

Qy         37 KGGTLNDFIKAADELGIARLAEEPNHTETAKKSVDTVNQFLSLTQTGIAISATKLEKFLQ 96
Db         61 QGNGVQDLVKAANDLGIEVWREERSNLDIAKTSFDTTQKILGFTDRGIVLFAPQLDNLLK 120

Qy         97 KHSTNKLAKGLDSVENIDRKLGKASNVLSTLSSFLGTALAGIELDSLIKKGDAAPDALAK 156
Db        121 KNP--KIGNTLGSASSISQNIGKANTVLGGIQSILGSVLSGVNLNELLQNKDPNQLELAK 178

Qy        157 ASIDLINEIIGNLSQSTQTIEAFSSQLAKLGSTISQAKGFSNIGNKLQNL-NFSKTNLGL 215
Db        179 AGLELTNELVGNIASSVQTVDAFAEQISKLGSHLQNVKGLGGLSNKLQNLPDLGKASLGL 238

Qy        216 EIITGLLSGISAGFALADKNASTGKKVAAGFELSNQVIGNVTKAISSYVLAQRVAAGLST 275
Db        239 DIISGLLSGASAGLILADKEASTEKKAAAGVEFANQIIGNVTKAVSSYILAQRVASGLSS 298

Qy        276 TGAVAALITSSIMLAISPLAFMNAADKFNHANALDEFAKQFRKFGYDGDHLLAEYQRGVG 335
Db        299 TGPVAALIASTVALAVSPLSFLNVADKFKQADLIKSYSERFQKLGYDGDRLLADFHRETG 358

Qy        336 TIEASLTTISTALGAVSAGVSAAAVGSAVGTPIALLVAGVTGLISGILEASKQAMFESVA 395
Db        359 TIDASVTTINTALAAISGGVGAASAGSLVGAPVALLVAGVTGLITTILEYSKQAMFEHVA 418

Qy        396 NRLQGKILEWEKQNGGQNYFDKGYDSRYAAYLANNLKFLSELNKELEAERVIAITQQRWD 455
Db        419 NKVHDRIVEWEKKH-NKNYFEQGYDSRHLADLQDNMKFLINLNKELQAERVVAITQQRWD 477

Qy        456 NNIGELAGITKLGERIKSGKAYADAFEDGKKVEAGSNITLDAKTGIIDISNSNGKKTQAL 515
Db        478 NQIGDLAAISRRTDKISSGKAYVDAFEEGQHQSYDSSVQLDNKNGIINISNTN-RKTQSV 536

Qy        516 HFTSPLLTAGTESRERLTNGKYSYINKLKFGRVKNWQVTDGEASSKLDFSKVIQRVA--- 572
Db        537 LFRTPLLTPGEENRERIQEGKNSYITKLHIQRVDSWTVTDGDASSSVDFTNVVQRIAVKF 596

Qy        573 -------ETEGTDEIGLIVNAKAGNDDIFVGQGKMNIDGGDGHDRVFYSKDGGFGNITVD 625
Db        597 DDAGNIIESKDTK---IIANLGAGNDNVFVGSSTTVIDGGDGHDRVHYSR-GEYGALVID 652

Qy        626 GTSATEAGSYTVNRKVARGDIYHEVVKRQETKVGKRTETIQYRDYELRKVGYGYQSTDNL 685
Db        653 ATAETEKGSYSVKRYVGDSKALHETIATHPTNVGNREEKIEYR-REDDRFHTGYTVTDSL 711

Qy        686 KSVEEVIGSQFNDVFKGSKFNDIFHSGEGDDLLDGGAGDDRLFGGKGNDRLSGDEGDDLL 745
Db        712 KSVEEIIGSQFNDIFKGSQFDDVFHGGNGVDTIDGNDGDDHLFGGAGDDVIDGGNGNNFL 771

Qy        746 DGGSGDDVLNGGAGNDVYIFRKGDGNDTLYDGTGNDKLAFADANISDIMIERTKEGIIVK 805
Db        772 VGGTGNDIISGGKDNDIYVHKTGDGNDSITDSGGQDKLAFSDVNLKDLTFKKVDSSLEI- 830

Qy        806 RNDHSGSINIPRWY----ITSNLQNYQSNKTDHKIEQLIGKDGSYITSDQIDKILQDKKD 861
Db        831 INQKGEKVRIGNWFLKNDLASTVANYKAT-NDRKIEEIIGKGGERITSKQVDKLI--KEG 887
```

```
Qy        862 GTVITSQELKKLADENKSQKLSASDIASSLNKLVGSMALFGTANSVSSN 910

Db        888 NNQISAKALSKVGNDYNTSK-DRQNVSNSLAKLISSVESFTSSSNFRNN 935

RESULT 20
Q8KWZ6
ID   Q8KWZ6      PRELIMINARY;      PRT;   987 AA.
AC   Q8KWZ6;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   AqxA.
GN   AQXA.
OS   Actinobacillus equuli.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC   Pasteurellaceae; Actinobacillus.
OX   NCBI_TaxID=718;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=90-570;
RX   MEDLINE=22031201; PubMed=12034544;
RA   Berthoud H., Frey J., Kuhnert P.;
RT   "Characterization of Aqx and its operon: the hemolytic RTX determinant
RT   of Actinobacillus equuli.";
RL   Vet. Microbiol. 87:159-174(2002).
DR   EMBL; AF381185; AAM45569.1; -.
DR   InterPro; IPR001343; Hemlysn_Ca_bind.
DR   InterPro; IPR003995; RtxA.
DR   Pfam; PF00353; hemolysinCabind; 3.
DR   Pfam; PF02382; RTX; 1.
DR   PRINTS; PR00313; CABNDNGRPT.
DR   PRINTS; PR01488; RTXTOXINA.
SQ   SEQUENCE   987 AA;  105871 MW;  35440030AB522E53 CRC64;

  Query Match             46.7%;  Score 2171;  DB 2;  Length 987;
  Best Local Similarity   48.1%;  Pred. No. 5e-89;
  Matches  468;  Conservative  163;  Mismatches  264;  Indels   78;  Gaps   23;

Qy          3 NINV--------IKSNIQAGLNSTKSGL-------KN----LYLAIPKDYDPQKGGTLND 43

Db         17 NLNLQFQQGAGKIKSGVQKGYQSSKQHLSQAGAATKNYANKLILYIPKEYDATAGNGLKG 76

Qy         44 FIKAADELGIARLAEEPNHTETAKKSVDTVNQFLSLTQTGIAISATKLEKFLQKHSTNKL 103

Db         77 LLDAAEALGIEVQREEKDGVATAYSAFGVGEQFIGLTERGLTLFLPQLDNFLKKHP--KF 134

Qy        104 AKGL-DSVENIDRKLGKASNVLSTLSSFLGTALAGIELDSLIKKGDAAPDALAKASIDLI 162

Db        135 SKSIGDSSEKVLSNVGNAKLILSGIQSVLGTTLAGIELDNLIKEGGSKTE-LAQAGINLV 193

Qy        163 NEIIGNLSQSTQTIEAFSSQLAKLGSTISQAKGFSNIGNKLQNLN---FSKTNLGLEIIT 219

Db        194 NELVGNIAKGAETIQAFSDQLAHLGSYIQNTKGLGGLGNKLQNISGSALNKAGLGFDIVS 253

Qy        220 GLLSGISAGFALADKNASTGKKVAAGFELSNQVIGNVTKAISSYVLAQRVAAGLSTTGAV 279

Db        254 SLLSGVSAGFALADPNASTGKKIAAGFELSNQVIGGITKAVSGYILAQRIASGLSTTGPA 313

Qy        280 AALITSSIMLAISPLAFMNAADKFNHANALDEFAKQFRKFGYDGDHLLAEYQRGVGTIEA 339

Db        314 IGLIASSISLAISPLAFLNVADRFKHANEINKFAERFKKFGYEGDSLLADFHHESGAIDA 373

Qy        340 SLTTISTALGAVSAGVSAAAVGSAVGTPIALLVAGVTGLISGILEASKQAMFESVANRLQ 399

Db        374 SITTINTALGAISAGVTAAGAASLVGAPVTLLVGGITGLISAILDLSKQAMFEHVATKLS 433

Qy        400 GKILEWEKQNGGQNYFDKGYDSRYAAYLANNLKFLSELNKELEAERVIAITQQRWDNNIG 459

Db        434 GKIEAWEKKY-GMNYFEKGYDARHAAFLEDNFAFFESLTKELRAERVISITQQQWDAQIG 492

Qy        460 ELAGITKLGERIKSGKAYADAFEDGKKVEAGSNI-TLDAKTGIIDISNSNGKKTQALHFT 518

Db        493 DLAGITRRGDKIQSGKAYVDVFKESKEDRNYDNIVTFDPTEGIIDISKTT-SKTQHLLFL 551

Qy        519 SPLLTAGTESRERLTNGKYSYINKLKFGRVKNWQVTDGEASSKLDFSKVIQRVAETEGTD 578

Db        552 NPLLTPGKENREREKKGKYEYVTKLIVDRKTKWQVTDGEASSTLDFTNVLQFIA--VDTD 609

Qy        579 EIGLI-------VNAK--AGNDDIFVGQGKMNIDGGDGHDRVFYSKDGGFGNI---TVDG 626

Db        610 RAGNVTESLEAKIEAKLGKGDDTVFVGWGSTDIDGSEGYDRAAYNRLDGEAHIYGLNIDA 669

Qy        627 TSATEAGSYTVNRKVARGDIYHEVVKRQETKVGKRTETIQYRDYELR-----KVGYGYQS 681

Db        670 QQETVAGSYTVNRTIGRGAAKHEVIKVHQATAGKRVDKIEYRQGESRFHNELKV------ 723

Qy        682 TDNLKSVEEVIGSQFNDVFKGSKFNDIFHSGEGDDLLDGGAGDDRLFGGKGNDRLSGDEG 741

Db        724 VDKLVNVEEIIGTKNQDIFKGSKFNDTFFGGDEKDIIYGNAGNDRLFGGNGNDEIDGGDG 783

Qy        742 DDLLDGGSGDDVLNGGAGNDVYIFRKGDGNDTLYDGTGNDKLAFADANISDIMIERTKEG 801

Db        784 NDFIDGGKGNDSLTGGYGNDIYVIHKEDGQDFITELGGDDRLVFTDTELSRLTFKRDKKV 843

Qy        802 IIVKRNDHSGSINIPRWYI----TSNLQNYQSN--KTDHKIEQLIGKDGSYITSDQIDKI 855

Db        844 LIIEDSAHNNQVRISDWFFKDQATLNKEYHAQSALDKDYRIEEIIDKNGTRITADQIDTI 903

Qy        856 LQDKKDGTVITSQELKKLADENKSQKLSA-----SDIASSLNKLVGSMALF--------G 902

Db        904 LNGKSEGVIEVSQ-LVKISDDYKA-KVAADNAHLSEIAEALGKIASSSASFSAVTGGAMG 961

Qy        903 TAN---SVSSNAL 912

Db        962 SANAFAAVSENAL 974

RESULT 21
Q8KWZ9
ID   Q8KWZ9      PRELIMINARY;      PRT;   987 AA.
AC   Q8KWZ9;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   AqxA.
GN   AQXA.
OS   Actinobacillus cf. equuli.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC   Pasteurellaceae; Actinobacillus.
OX   NCBI_TaxID=182091;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CCUG 19799;
RX   MEDLINE=22031201; PubMed=12034544;
RA   Berthoud H., Frey J., Kuhnert P.;
RT   "Characterization of Aqx and its operon: the hemolytic RTX determinant
RT   of Actinobacillus equuli.";
RL   Vet. Microbiol. 87:159-174(2002).
DR   EMBL; AF381184; AAM45566.1; -.
DR   InterPro; IPR001343; Hemlysn_Ca_bind.
DR   InterPro; IPR003995; RtxA.
DR   Pfam; PF00353; hemolysinCabind; 1.
DR   Pfam; PF02382; RTX; 1.
DR   PRINTS; PR00313; CABNDNGRPT.
DR   PRINTS; PR01488; RTXTOXINA.
DR   PROSITE; PS00330; HEMOLYSIN_CALCIUM; 1.
SQ   SEQUENCE   987 AA;  105662 MW;  CC34CD62BF9B745D CRC64;

  Query Match             46.4%;  Score 2155;  DB 2;  Length 987;
  Best Local Similarity   47.5%;  Pred. No. 2.6e-88;
  Matches  463;  Conservative  167;  Mismatches  266;  Indels   78;  Gaps   23;

Qy          2 SNINV--------IKSNIQAGLNSTKSGL-------KN----LYLAIPKDYDPQKGGTLN 42
```

```
Db         16 SNLNLQFQQGAGKIKSGVQKGYQSSKQHLSQAGAATKNYANKLILYIPKEYDATAGNGLK 75

Qy         43 DFIKAADELGIARLAEEPNHTETAKKSVDTVNQFLSLTQTGIAISATKLEKFLQKHSTNK 102
Db         76 GLLDAAEALGIEVQREEKDGVATAYSAFGVGEQFIGLTERGLTLFLPQLDNFLKKHP--K 133

Qy        103 LAKGL-DSVENIDRKLGKASNVLSTLSSFLGTALAGIELDSLIKKGDAAPDALAKASIDL 161
Db        134 FSKSIGDSSEKVLGNVGNAKLILSGIQSVLGTTLAGIELDNLIKEGGSKTE-LAQAGINL 192

Qy        162 INEIIGNLSQSTQTIEAFSSQLAKLGSTISQAKGFSNIGNKLQNLN---FSKTNLGLEII 218
Db        193 VNELVGNIAKGAETIQAFSDQLAHLGSYIQNTKGLGGLGNKLQNISGSALNKAGLGFDIV 252

Qy        219 TGLLSGISAGFALADKNASTGKKVAAGFELSNQVIGNVTKAISSYVLAQRVAAGLSTTGA 278
Db        253 SSLLSGVSAGFALADPNASTGKKIAAGFELSNQVIGGITKAVSGYILAQRIASGLSTTGP 312

Qy        279 VAALITSSIMLAISPLAFMNAADKFNHANALDEFAKQFRKFGYDGDHLLAEYQRGVGTIE 338
Db        313 AIGLIASSISLAISPLAFLNVADKFNYAKEINKFAERFKKFGYEGDSLLADFHHESGAID 372

Qy        339 ASLTTISTALGAVSAGVSAAAVGSAVGTPIALLVAGVTGLISGILEASKQAMFESVANRL 398
Db        373 AAITTINTALGAISAGVTAAGAASLVGAPVTLLVGGVTGLISAILDLSKQAMFEHVATKL 432

Qy        399 QGKILEWEKQNGGQNYFDKGYDSRYAAYLANNLKFLSELNKELEAERVIAITQQRWDNNI 458
Db        433 SGKIEAWEKKY-GMNYFEKGYDARHAAFLEDNFAFFESLTKELRAERVISITQQQWDTQI 491

Qy        459 GELAGITKLGERIKSGKAYADAFEDGKKVEAGSNI-TLDAKTGIIDISNSNGKKTQALHF 517
Db        492 GDLAGITRRGDKIQSGKAYVDVFKETKKDLSYDNIVTFDPTEGIIDISKTT-SKTQHLLF 550

Qy        518 TSPLLTAGTESRERLTNGKYSYINKLKFGRVKNWQVTDGEASSKLDFSKVIQRVAETEGT 577
Db        551 LNPLLTPGKENREREKKGKYEYVTKLIVDRKTKWQVTDGEASSTLDFTNVLQFIA--VDT 608

Qy        578 DEIGLI-------VNAK--AGNDDIFVGQGKMNIDGGDGHDRVFYSKDGGFGNI---TVD 625
Db        609 DRAGNVTESLEAKIEAKLGKGDDTVFVGWGSTDIDGGEGYDRAAYNRLDGEAHIYGLNID 668

Qy        626 GTSATEAGSYTVNRKVARGDIYHEVVKRQETKVGKRTETIQYRDYELR-----KVGYGYQ 680
Db        669 AQQETVAGSYTVNRTIAGGAAKHEVIKVHQATAGKRVDKIEYRQGEQRYHQDLKV----- 723

Qy        681 STDNLKSVEEVIGSQFNDVFKGSKFNDIFHSGEGDDLLDGGAGDDRLFGGKGNDRLSGDE 740
Db        724 -VDKLLNIEEIIGTKNRDNFKGSKFNDTFFGGDEKDIIDGNAGNDRLFGGDGGDEIDGGD 782

Qy        741 GDDLLDGGSGDDVLNGGAGNDVYIFRKGDGNDTLYDGTGNDKLAFADANISDIMIERTKE 800
Db        783 GNDFIDSGKGDDIISGGHGNDIYVIHKDGAKDYITELGGDDRLVFTDTELSRLTFKRDNR 842

Qy        801 GIIVKRNDHSGSINIPRWYITSNL---QNYQSNKT---DHKIEQLIGKDGSYITSDQIDK 854
Db        843 TLVIEDSANNNHVQVSDWFLKDQAKINKEYHAQSVLDKDYRIEEIIDKNGTRITADQIDK 902

Qy        855 ILQDKKDGTVITSQELKKLADENKSQKLSA-----SDIASSLNKLVGSMALF-------- 901
Db        903 ILNGKSEGVIEVSQ-LVKISDDYKA-KVAADNAHLSEIAEALGKIASSSASFSAVTGGAM 960

Qy        902 GTAN---SVSSNAL 912
Db        961 GSANAFAAVSENAL 974
```

RESULT 22
Q9RCG8

```
ID   Q9RCG8      PRELIMINARY;      PRT;  1049 AA.
AC   Q9RCG8;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   PaxA.
GN   PAXA.
OS   Pasteurella aerogenes.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC   Pasteurellaceae; Pasteurella.
OX   NCBI_TaxID=749;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=field strain JF1319;
RA   Heyberger-Meyer B., Frey J., Nicolet J., Kuhnert P.;
RT   "Identification and characterization of a new RTX toxin operon (pax)
RT   in Pasteurella aerogenes.";
RL   Submitted (AUG-1996) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U66588; AAF15370.1; -.
DR   InterPro; IPR001343; Hemlysn_Ca_bind.
DR   InterPro; IPR003995; RtxA.
DR   Pfam; PF00353; hemolysinCabind; 6.
DR   Pfam; PF02382; RTX; 1.
DR   PRINTS; PR00313; CABNDNGRPT.
DR   PRINTS; PR01488; RTXTOXINA.
DR   PROSITE; PS00330; HEMOLYSIN_CALCIUM; 1.
SQ   SEQUENCE   1049 AA;  112308 MW;  9FA5070E48CC3127 CRC64;

  Query Match             43.1%;  Score 2004.5;  DB 2;  Length 1049;
  Best Local Similarity   43.7%;  Pred. No. 1.5e-81;
  Matches  439;  Conservative  158;  Mismatches  288;  Indels  119;  Gaps   20;

Qy          6 VIKSNIQAGLNSTKSGLK---------------------NLYLAIPKDYDPQKGGTLND 43
Db         17 IAKQQAKKGIDVTKNGLQYGVSQVKLQALAAGKSIQKYGNKLVLVIPKDYDVNTGNGFFD 76

Qy         44 FIKAADELGIARLAEEPNHTETAKKSVDTVNQFLSLTQTGIAISATKLEKFLQKHSTNKL 103
Db         77 LAKAAEELGIQVKYIDRNDLEIAHKSLGVTDQFLGLTERGLTLFAPQLDKFLQQHSKISN 136

Qy        104 AKGLDSVENIDRKLGKASNVLSTLSSFLGTALAGIELDSLIKKGDAAPDALAKASIDLIN 163
Db        137 VVGSSTGDTVN-KLAKSQAIISGVQSVLGSVLAGINLNEAIISGGSELE-LAKAGVDLAS 194

Qy        164 EIIGNLSQSTQTIEAFSSQLAKLGSTISQAKGFSNIGNKLQNLN---FSKTNLGLEIITG 220
Db        195 ELVGNIAKGTATIEAFSEQIQNFGKLVQNAKGLGGVGQQLQHISGSALSKTGLGLDIISS 254

Qy        221 LLSGISAGFALADKNASTGKKVAAGFELSNQVIGNVTKAISSYVLAQRVAAGLSTTGAVA 280
Db        255 LLSGVTASFTLADKNASTSTKVAAGFELSNQVIGGITKAVSSYILAQRLAAGLSTTGPAA 314

Qy        281 ALITSSIMLAISPLAFMNAADKFNHANALDEFAKQFRKFGYDGDHLLAEYQRGVGTIEAS 340
Db        315 ALIASSISLAISPLSFLRVADNFNRSKDIREFAERFKKLGYEGDKLLSDFYHEAGTIDAS 374

Qy        341 LTTISTALGAVSAGVSAAAVGSAVGTPIALLVAGVTGLISGILEASKQAMFESVANRLQG 400
Db        375 ITTISTALSAIAAGTAAASAGALVGAPITLLVTGITGLISGILEFSKQPMLEHVASKLGT 434

Qy        401 KILEWEKQNGGQNYFDKGYDSRYAAYLANNLKFLSELNKELEAERVIAITQQRWDNNIGE 460
Db        435 KIEEWERKY-GKNYFENGYDARHKAFLEDSLSLLSSFNKQYETERAVLITQQRWDEYIGE 493

Qy        461 LAGITKLGERIKSGKAYADAFEDG----KKVEAGSNITLDAKTGIIDISNSNGKKTQALH 516
Db        494 LAGVTGKGDKISSGKAYVDYFEEGKLLAKKPDDFNRVILDPKKGKIDISNS--QTSTLLK 551

Qy        517 FTSPLLTAGTESRERLTNGKYSYINKLKFGRVKNWQVTD-GEASSKLDFSKVIQR----- 570
Db        552 FVTPLLTPGTESRKRTQTGKYEYVTKLDVNGINQWEVNGVKEKGAVYDFTNLIQHVHISS 611

Qy        571 -VAETEGTDEIGLIVNAKAGNDDIFVGQGKMNIDGGDGHDRVFYSKDGGFGNITVDGTSA 629
Db        612 SVARGEEYREVRLVSRLGKGNDKVFLASGSAEIHAGDGHDVVYYDKT-DTGLLMVDGTQA 670

Qy        630 TEAGSYTVNRKVARG-DIYHEVVKRQETKVGKRTETIQYRDYELRKVG-YGYQSTDNLKS 687
```

```
               |: | ||| |:::    | |||| |:: || | ||::||| || : |     :: ||| |
Db        671 TKQGDYTVTRELSGATQILREVVKNQKSSVGSRQETVEYRDNELAQSGNSNLKAKDNLYS 730

Qy        688 VEEVIGSQFNDVFKGSKFNDIFHSGEGDDLLD---------------------------G 720
              |||:|||    | |||||| ||||  :|||||:                           |
Db        731 VEEIIGSNHRDEFKGSKFRDIFHGADGDDLLNGNDGDDILYGDKGNDELRGDNGNDQLYG 790

Qy        721 GAGDDRLFGGKGNDRLSGDEGD---------------------------DLLDGGSGDD 752
              | |:|:|||| ||: ||| :||                           | |||| |||
Db        791 GEGNDKLFGGNGNNYLSGGDGDDELQVLGNGFNVLRGGKGNDKLYGGAGSDFLDGGEGDD 850

Qy        753 VLNGGAGNDVYIFRKGDGNDTLYD---GTGNDKLAFADANISDIMIERTKEGIIVKRNDH 809
                | ||| |||  |::|     || |:||     :::|:   :::|:      :: | :||
Db        851 YLAGGEGNDFYVYRSTSGNHTIYDQGKSSDSDTLYLSDLTFDRLLVEKVDNNLVFKPSDH 910

Qy        810 S---GSINIPRWYITSNLQNYQSNKTDHKIEQLIGKDGSYITSDQIDKILQDKKDGTVIT 866
              :    ||: |  |: | :  |         ||:||:: |:|  :|||      |:   :|| |
Db        911 NSNRGSLTIKDWFKTGHGYN-------HKLEQIVDKNGRKLTSDN----LETHFNGTPKT 959

Qy        867 SQELKKLADENKSQKLSASDIASSLNKLVGSMALFGTANSVSSN 910
              :  |      ||:::  : | : : | |:: |     ||| |   ::|
Db        960 N--LLGYTAENQNES-NLSSLKTELGKIISSAGNFGLAKQGNNN 1000
```

RESULT 23
Q93NP0
ID   Q93NP0   PRELIMINARY;   PRT; 1052 AA.
AC   Q93NP0;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   RTX-toxin IIIA.
OS   Actinobacillus pleuropneumoniae (Haemophilus pleuropneumoniae).
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC   Pasteurellaceae; Actinobacillus.
OX   NCBI_TaxID=715;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Shin S.J., Park J.Y., Choi I.S., Shin N.R., Yoo H.S.;
RT   "Cloning and Sequencing of Apx IIIA from Actinobacillus
RT   pleuropneumoniae.";
RL   Submitted (MAR-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF363363; AAK50053.1; -.
DR   InterPro; IPR001343; Hemlysn_Ca_bind.
DR   InterPro; IPR003995; RtxA.
DR   Pfam; PF00353; hemolysinCabind; 6.
DR   Pfam; PF02382; RTX; 1.
DR   PRINTS; PR00313; CABNDNGRPT.
DR   PRINTS; PR01488; RTXTOXINA.
DR   PROSITE; PS00330; HEMOLYSIN_CALCIUM; 1.
SQ   SEQUENCE   1052 AA; 112862 MW; CA0E160F02101F42 CRC64;

```
  Query Match             42.1%;  Score 1956.5;  DB 2;  Length 1052;
  Best Local Similarity   42.6%;  Pred. No. 2.1e-79;
  Matches  425;  Conservative  167;  Mismatches  286;  Indels  119;  Gaps   19;

Qy          8 KSNIQAGLNSTKSGLK---------------------NLYLAIPKDYDPQKGGTLNDFI 45
              |   :  | :  ||:||:                    | | |||:||  |     |  :
Db         19 KRQAKKGYDVTKNGLQYGVSQAKLQALAAGKAVQKYGNKLVLVIPKEYDGSVGNGFFDLV 78

Qy         46 KAADELGIARLAEEPNHTETAKKSVDTVNQFLSLTQTGIAISATKLEKFLQKHSTNKLAK 105
              |||:||||        |  |  |  ||: |  :|||  ||: |: :|| :|::||||||
Db         79 KAAEELGIQVKYVNRNELEVAHKSLGTADQFLGLTERGLTLSAPQLDQFLQKHSKISNVV 138

Qy        106 GLDSVENIDRKLGKASNVLSTLSSFLGTALAGIELDSLIKKGDAAPDALAKASIDLINEI 165
              |  : : :  || |:  ::| : | ||| |||| |:  |   |  :  : ||:| : | :|:
Db        139 GSSTGDAVS-KLAKSQTIISGIQSVLGTVLAGINLNEAIISGGSELE-LAEAGVSLASEL 196

Qy        166 IGNLSQSTQTIEAFSSQLAKLGSTISQAKGFSNIGNKLQNLN---FSKTNLGLEIITGLL 222
              : |::: | ||:||::|:   |  :  |||   :| :|||::     ||| |||:||: ||
Db        197 VSNIAKGTTTIDAFTTQIQNFGKLVENAKGLGGVGRQLQNISGSALSKTGLGLDIISSLL 256

Qy        223 SGISAGFALADKNASTGKKVAAGFELSNQVIGNVTKAISSYVLAQRVAAGLSTTGAVAAL 282
              ||::| ||||:|||||  |||||||||||||| :|||:|||:||||:||||||||  |||
Db        257 SGVTASFALANKNASTSTKVAAGFELSNQVIGGITKAVSSYILAQRLAAGLSTTGPAAAL 316

Qy        283 ITSSIMLAISPLAFMNAADKFNHANALDEFAKQFRKFGYDGDHLLAEYQRGVGTIEASLT 342
              | ||| ||||||||:  || ||  :  : |||::|:| ||||| ||:|:      |||:||:|
Db        317 IASSISLAISPLAFLRVADNFNRSKEIGEFAERFKKLGYDGDKLLSEFYHEAGTIDASIT 376

Qy        343 TISTALGAVSAGVSAAAVGSAVGTPIALLVAGVTGLISGILEASKQAMFESVANRLQGKI 402
              ||||| |::|| :||: |: || || ||| |:||||||||| ||| |  : ||::: ||
Db        377 TISTALSAIAAGTAAASAGALVGAPITLLVTGITGLISGILEFSKQPMLDHVASKIGNKI 436

Qy        403 LEWEKQNGGQNYFDKGYDSRYAAYLANNLKFLSELNKELEAERVIAITQQRWDNNIGELA 462
              ||||:   |:|||: |||:|: |:|  ||  ||: |||  : |||||||  |||||
Db        437 DEWEKKY-GKNYFENGYDARHKAFLEDSFSLLSSFNKQYETERAVLITQQRWDEYIGELA 495

Qy        463 GITKLGERIKSGKAYADAFEDGKKVEAG----SNITLDAKTGIIDISNSNGKKTQALHFT 518
              ||| |:::  |||| |  |::|| :|         |  :  |  ||||||  :  :  | |
Db        496 GITGKGDKLSSGKAYVDYFQEGKLLEKKPDDFSKVVFDPTKGEIDISNS--QTSTLLKFV 553

Qy        519 SPLLTAGTESRERLTNGKYSYINKLKFGRVKNWQVTDG--EASSKLDFSKVIQR------ 570
              :|||| ||||||| ||| || ||      | | :|  :  :  |:: :||
Db        554 TPLLTPGTESRERTQTGKYEYITKLVVKGKDKW-VVNGVKDKGAVYDYTNLIQHAHISSS 612

Qy        571 VAETEGTDEIGLIVNAKAGNDDIFVGQGKMNIDGGDGHDRVFYSKDGGFGNITVDGTSAT 630
              || |  |: |: :   ||| :|:  |  |  |:||| |:| |     | : :||| ||
Db        613 VARGEEYREVRLVSHLGNGNDKVFLAAGSAEIHAGEGHDVVYYDKT-DTGLLVIDGTKAT 671

Qy        631 EAGSYTVNRKVARG-DIYHEVVKRQETKVGKRTETIQYRDYELRKVG-YGYQSTDNLKSV 688
              | | |:| |:::    | ||:| |::| ||| ||:: ||||||| :  |    :: | | ||
Db        672 EQGRYSVTRELSGATKILREVIKNQKSAVGKREETLEYRDYELTQSGNSNLKAHDELHSV 731

Qy        689 EEVIGSQFNDVFKGSKFNDIFHSGEGDDLLDGGAGDDRLFGGKGNDRLSGDEGD------ 742
              ||:|||  | |||||| ||||  :|||||:| ||| |: |||| || ||  :
Db        732 EEIIGSNQRDEFKGSKFRDIFHGADGDDLLNGNDGDDILYGDKGNDELRGDNGNDQLYGG 791

Qy        743 ------------------------------------------------DLLDGGSGDDV 753
                                                              |||||| |:|
Db        792 EGNDKLLGGNGNNYLSGGDGNDELQVLGNGFNVLRAGKGDDKLYGSSGSDLLDGGEGNDY 851

Qy        754 LNGGAGNDVYIFRKGDGNDTLYD---GTGNDKLAFADANISDIMIERTKEGIIVKRNDHS 810
              | || |:| |::|     || |:||     :  |||  :| :   :::|:  : :::: |: |
Db        852 LEGGDGSDFYVYRSTSGNHTIYDQGKSSDLDKLYLSDFSFDRLLVEKVDDNLVLRSNESS 911

Qy        811 GS---INIPRWYITSNLQNYQSNKTDHKIEQLIGKDGSYITSDQIDKILQDKKDGTVITS 867
              :    :   |:         : || :||||:: |:| :  :|:: :  :       :  :
Db        912 HNNRVLTIKDWF-------KEGNKYNHKIEQIVDKNGRKLTAENLGTYFKNAPKADNLLN 964

Qy        868 QELKKLADENKSQKLSASDIASSLNKLVGSMALFGTA 904
                  |:   |:|:|   : |  : : |:|:: :    |||
Db        965 YATKE--DQNES---NLSSLKTELSKIITNAGNFGVA 996
```

RESULT 24
Q8FE01
ID   Q8FE01   PRELIMINARY;   PRT; 1024 AA.
AC   Q8FE01;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Hemolysin A.
GN   HLYA OR C3570.
OS   Escherichia coli O6.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Escherichia.
OX   NCBI_TaxID=217992;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=O6:H1 / CFT073 / ATCC 700928;
RX   MEDLINE=22388234; PubMed=12471157;
RA   Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,
RA   Rasko D., Buckles E.L., Liou S.-R., Boutin A., Hackett J., Stroud D.,
RA   Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,
RA   Mobley H.L.T., Donnenberg M.S., Blattner F.R.;

```
RT   "Extensive mosaic structure revealed by the complete genome sequence
RT   of uropathogenic Escherichia coli.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024(2002).
DR   EMBL; AE016766; AAN82018.1; -.
KW   Complete proteome.
SQ   SEQUENCE   1024 AA;  110246 MW;  9DEBA50C23B5A49C CRC64;

  Query Match             39.8%;  Score 1848.5;  DB 16;  Length 1024;
  Best Local Similarity   41.8%;  Pred. No. 1.4e-74;
  Matches  428;  Conservative  153;  Mismatches  310;  Indels  132;  Gaps   25;

Qy          2 SNINVIKSNIQAGLNSTKSGLK-----------NLYLAIPKDYDPQKGGTLNDFIKAADE 50
Db         19 SSANKLHS---AG-QSTKDALKKAAEQTRNAGNRLILLIPKDYKGQ-GSSLNDLVRTADE 73

Qy         51 LGIARLAEEPNHTETAKKSVDTVNQFLSLTQTGIAISATKLEKFLQKH--STNKLAKGLD 108
Db         74 LGIEVQYDEKNGTAITKQVFGTAEKLIGLTERGVTIFAPQLDKLLQKYQKAGNKLG---G 130

Qy        109 SVENIDRKLGKASNVLSTLSSFLGTALAGIELDSLIKK----GDAAPDALAKASIDLINE 164
Db        131 SAENIGDNLGKAGSVLSTFQNFLGTALSSMKIDELIKRQKSGSNVSSSELAKASIELINQ 190

Qy        165 IIGNLSQSTQTIEAFSSQLAKLGSTISQAKGFSNIGNKLQNL-NFSKTNLGLEIITGLLS 223
Db        191 LVDTAASINNNVNSFSQQLNKLGSVLSNTKHLTGVGNKLQNLPNLDNIGAGLDTVSGILS 250

Qy        224 GISAGFALADKNASTGKKVAAGFELSNQVIGNVTKAISSYVLAQRVAAGLSTTGAVAALI 283
Db        251 AISASFILSNADADTGTKAAAGVELTTKVLGNVGKGISQYIIAQRAAQGLSTSAAAAGLI 310

Qy        284 TSSIMLAISPLAFMNAADKFNHANALDEFAKQFRKFGYDGDHLLAEYQRGVGTIEASLTT 343
Db        311 ASVVTLAISPLSFLSIADKFKRANKIEEYSQRFKKLGYDGDSLLAAFHKETGAIDASLTT 370

Qy        344 ISTALGAVSAGVSAAAVGSAVGTPIALLVAGVTGLISGILEASKQAMFESVANRLQGKIL 403
Db        371 ISTVLASVSSGISAAATTSLVGAPVSALVGAVTGIISGILEASKQAMFEHVASKMADVIA 430

Qy        404 EWEKQNGGQNYFDKGYDSRYAAYLANNLKFLSELNKELEAERVIAITQQRWDNNIGELAG 463
Db        431 EWEKKH-GKNYFENGYDARHAAFLEDNFKILSQYNKEYSVERSVLITQQHWDTLIGELAG 489

Qy        464 ITKLGERIKSGKAYADAFEDGKKVEAG----SNITLDAKTGIIDISNSNGKKTQALHFTS 519
Db        490 VTRNGDKTLSGKSYIDYYEEGKRLEKKPDEFQKQVFDPLKGNIDLSDS--KSSTLLKFVT 547

Qy        520 PLLTAGTESRERLTNGKYSYINKLKFGRVKNWQVTD-GEASSKLDFSKVIQRVAETEGT- 577
Db        548 PLLTPGEEIRERRQSGKYEYITELLVKGVDKWTVKGVQDKGSVYDYSNLIQHASVGNNQY 607

Qy        578 DEIGLIVNAKAGNDDIFVGQGKMNIDGGDGHDRVFYSK-DGGFGNITVDGTSATEAGSYT 636
Db        608 REIRIESHLGDGDDKVFLSAGSANIYAGKGHDVVYYDKTDTGY--LTIDGTKATEAGNYT 665

Qy        637 VNRKVARGD--IYHEVVKRQETKVGKRTETIQYRDYELRKV-GYGYQSTDNLKSVEEVIG 693
Db        666 VTR-VLGGDVKILQEVVKEQEVSVGKRTEKTQYRSYEFTHINGKNLTETDNLYSVEELIG 724

Qy        694 SQFNDVFKGSKFNDIFHSGEGDDLLDGGAGDDRLF------------------------- 728
Db        725 TTRADKFFGSKFTDIFHGADGDDHIEGNDGNDRLYGDKGNDTLRGGNGDDQLYGGDGNDK 784

Qy        729 -------------------------------GGKGNDRLSGDEGDDLLDGGSGDDVLNG 756
Db        785 LIGGTGNNYLNGGDGDDELQVQGNSLAKNVLSGGKGNDKLYGSEGADLLDGGEGNDLLKG 844

Qy        757 GAGNDVYIFRKGDGNDTLYDGTG-NDKLAFADANISDIMIERTKEGIIVKRND------- 808
Db        845 GYGNDIYRYLSGYGHHIIDDDGGKDDKLSLADIDFRDVAFRREGNDLIMYKAEGNVLSIG 904

Qy        809 HSGSINIPRWYITSNLQNYQSNKTDHKIEQLIGKDGSYITSDQIDKILQDKK-------- 860
Db        905 HKNGITFRNWF-----EKESGDISNHQIEQIFDKDGRVITPDSLKKALEYQQSNNKASYV 959

Qy        861 ---DGTVITSQ--------ELKKLADENKSQKLSASDIASSLNKLVGSMALFGTA-NSVS 908
Db        960 YGNDALAYGSQDNLNPLINEISKIISAAGNFDVKEERAAASLLQLSGNASDFSYGRNSIT 1019

Qy        909 SNA 911
Db       1020 LTA 1022


RESULT 25
Q8GA40
ID   Q8GA40      PRELIMINARY;      PRT;  1024 AA.
AC   Q8GA40;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Hemolysin A.
GN   HLYA.
OS   Escherichia coli.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Escherichia.
OX   NCBI_TaxID=562;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=536;
RX   MEDLINE=22267134; PubMed=12379716;
RA   Dobrindt U., Blum-Oehler G., Nagy G., Schneider G., Johann A.,
RA   Gottschalk G., Hacker J.;
RT   "Genetic structure and distribution of four pathogenicity islands (PAI
RT   I536 to PAI IV536) of uropathogenic Escherichia coli strain 536.";
RL   Infect. Immun. 70:6365-6372(2002).
DR   EMBL; AJ488511; CAD33759.1; -.
SQ   SEQUENCE   1024 AA;  110314 MW;  9771F0A61ECF6EEB CRC64;

  Query Match             39.7%;  Score 1845.5;  DB 2;  Length 1024;
  Best Local Similarity   42.1%;  Pred. No. 1.9e-74;
  Matches  424;  Conservative  153;  Mismatches  303;  Indels  128;  Gaps   25;

Qy         17 STKSGLK-----------NLYLAIPKDYDPQKGGTLNDFIKAADELGIARLAEEPNHTET 65
Db         30 STKDALKKAAEQNRNAGNRLILLIPKDYKGQ-GSSLNDLVRTADELGIEVQYDEKNGTAI 88

Qy         66 AKKSVDTVNQFLSLTQTGIAISATKLEKFLQKH--STNKLAKGLDSVENIDRKLGKASNV 123
Db         89 TKQVFGTAEKLIGLTERGVTIFAPQLDKLLQKYQKAGNKLG---GSAENIGDNLGKAGSV 145

Qy        124 LSTLSSFLGTALAGIELDSLIKK----GDAAPDALAKASIDLINEIIGNLSQSTQTIEAF 179
Db        146 LSTFQNFLGTALSSMKIDELIKKQKSGSNVSSSELAKASIELINQLVDTAASINNNVNSF 205

Qy        180 SSQLAKLGSTISQAKGFSNIGNKLQNL-NFSKTNLGLEIITGLLSGISAGFALADKNAST 238
Db        206 SQQLNKLGSVLSNTKHLNGVGNKLQNLPNLDNIGAGLDTVSGILSVISASFILSNADADT 265

Qy        239 GKKVAAGFELSNQVIGNVTKAISSYVLAQRVAAGLSTTGAVAALITSSIMLAISPLAFMN 298
Db        266 GTKAAAGVELTTKVLGNVGKGISQYIIAQRAAQGLSTSAAAAGLIASAVTLAISPLSFLS 325

Qy        299 AADKFNHANALDEFAKQFRKFGYDGDHLLAEYQRGVGTIEASLTTISTALGAVSAGVSAA 358
Db        326 IADKFKRANKIEEYSQRFKKLGYDGDSLLAAFHKETGAIDASLTTISTVLASVSSGISAA 385

Qy        359 AVGSAVGTPIALLVAGVTGLISGILEASKQAMFESVANRLQGKILEWEKQNGGQNYFDKG 418
Db        386 ATTSLVGAPVSALVGAVTGIISGILEASKQAMFEHVASKMADVIAEWEKKH-GKNYFENG 444

Qy        419 YDSRYAAYLANNLKFLSELNKELEAERVIAITQQRWDNNIGELAGITKLGERIKSGKAYA 478
Db        445 YDARHAAFLEDNFKILSQYNKEYSVERSVLITQQHWDTLIGELAGVTRNGDKTLSGKSYI 504

Qy        479 DAFEDGKKVEAG----SNITLDAKTGIIDISNSNGKKTQALHFTSPLLTAGTESRERLTN 534
```

```
               | :|:||::|              |    | ||:|:|  | :   |  | :|||  |  |  |||    :
Db       505 DYYEEGKRLEKKPDEFQKQVFDPLKGNIDLSDS--KSSTLLKFVTPLLTPGEEIRERRQS 562

Qy       535 GKYSYINKLKFGRVKNWQVTD-GEASSKLDFSKVIQRVAETEGT-DEIGLIVNAKAGNDD 592
             ||| || :|      |   |  |        :   |   |:| :||   :          ||   :    :    |:|
Db       563 GKYEYITELLVKGVDKWTVKGVQDKGSVYDYSNLIQHASVGNNQYREIRIESHLGDGDDK 622

Qy       593 IFVGQGKMNIDGGDGHDRVFYSK-DGGFGNITVDGTSATEAGSYTVNRKVARGD--IYHE 649
             :|:    |    ||    |   |||   |:|  |  |  |   |:   :|:|||  ||||:|||  |   ||   :   |
Db       623 VFLAAGSANIYAGKGHDVVYYDKTDTGY--LTIDGTKATEAGNYTVTR-VLGGDVKVLQE 679

Qy       650 VVKRQETKVGKRTETIQYRDYELRKV-GYGYQSTDNLKSVEEVIGSQFNDVFKGSKFNDI 708
             |||  ||    |||||    |||  ||       :   |          |||  ||||:||:   |  |||| ||
Db       680 VVKEQEVSVGKRTEKTQYRSYEFTHINGTDLTETDNLYSVEELIGTNRADKFFGSKFTDI 739

Qy       709 FHSGEGDDLLDGGAGDDRLF---------------------------------------- 728
             ||  :|||  ::|   |:|||:
Db       740 FHGADGDDHIEGNDGNDRLYGDKGNDTLRGGNGDDQLYGGDGNDKLTGGVGNNYLNGGDG 799

Qy       729 -----------------GGKGNDRLSGDEGDDLLDGGSGDDVLNGGAGNDVYIFRKGDGN 771
                              ||||||:|  |  ||  ||||||  |:|:|  || |||:|  :   |  |:
Db       800 DDELQVQGNSLAKNVLSGGKGNDKLYGSEGADLLDGGEGNDLLKGGYGNDIYRYLSGYGH 859

Qy       772 DTLYDGTG-NDKLAFADANISDIMIERTKEGIIVKRND-------HSGSINIPRWYITSN 823
               :  |    |  :|||:  ||  :   |:   :|      :|: :  :            |      |       |:
Db       860 HIIDDDGGKDDKLSLADIDFRDVAFKREGNDLIMYKAEGNVLSIGHKNGITFRNWF---- 915

Qy       824 LQNYQSNKTDHKIEQLIGKDGSYITSDQIDKIL---QDKKDGTVITSQELKKLADENK-- 878
               :      :  ::|:|||:  |||   ||  |  :  |      |              :     :          || :
Db       916 -EKESGDISNHQIEQIFDKDGRVITPDSLKKAFEYQQSNNQANYVYGEYASTYADLDNLN 974

Qy       879 ------SQKLSAS---DI-----ASSLNKLVGSMALFGTA-NSVSSNA 911
                   |: :||:     |:        |:|| :|  |: :  |     ||:: |
Db       975 PLINEISKIISAAGNFDVKEERSAASLLQLSGNASDFSYGRNSITLTA 1022
```

RESULT 26
Q8G9Z4
ID   Q8G9Z4   PRELIMINARY;   PRT; 1024 AA.
AC   Q8G9Z4;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   HlyA protein.
GN   HLYA.
OS   Escherichia coli.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Escherichia.
OX   NCBI_TaxID=562;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=536;
RX   MEDLINE=22267134; PubMed=12379716;
RA   Dobrindt U., Blum-Oehler G., Nagy G., Schneider G., Johann A.,
RA   Gottschalk G., Hacker J.;
RT   "Genetic structure and distribution of four pathogenicity islands (PAI
RT   I536 to PAI IV536) of uropathogenic Escherichia coli strain 536.";
RL   Infect. Immun. 70:6365-6372(2002).
DR   EMBL; AJ494981; CAD42039.1; -.
SQ   SEQUENCE   1024 AA; 110365 MW; 3F3C284A45FEFEB3 CRC64;

```
  Query Match             39.5%;  Score 1836.5;  DB 2;  Length 1024;
  Best Local Similarity   41.7%;  Pred. No. 4.7e-74;
  Matches  420;  Conservative  156;  Mismatches  313;  Indels  117;  Gaps   24;

Qy         8 KSNIQAGLNSTKSGLKNLYLAIPKDYDPQKGGTLNDFIKAADELGIARLAEEPNHTETAK 67
             |   ::       |::     | | |||||  | | :|||  :: ||||||       :| |  |    |
Db        32 KDALKKAAEQTRNAGNRLILLIPKDYKXQ-GSSLNDLVRTADELGIEVQYDEKNGTAITK 90

Qy        68 KSVDTVNQFLSLTQTGIAISATKLEKFLQKH--STNKLAKGLDSVENIDRKLGKASNVLS 125
             :      |  :  : ||: |: |  | :|:| |||:   :  |||          | |||      |||| :|||
Db        91 QVFGTAEKLIGLTERGVTIFAPQLDKLLQKYQKAGNKLG---GSAENIGDNLGKAGSVLS 147

Qy       126 TLSSFLGTALAGIELDSLIKK----GDAAPDALAKASIDLINEIIGNLSQSTQTIEAFSS 181
             |  :||||||: :::| ||||         |: :     ||||||:|||:::   :       : :||
Db       148 TFQNFLGTALSSMKIDELIKKQKSGGNVSSSELAKASIELINQLVDTAASLNNNVNSFSQ 207

Qy       182 QLAKLGSTISQAKGFSNIGNKLQNL-NFSKTNLGLEIITGLLSGISAGFALADKNASTGK 240
             || ||||  :|      |   :   :||||||  |       ||:  ::|:|| |||  |  |::  :| ||
Db       208 QLNKLGSVLSNTKHLNGVGNKLQNLPNLDNIGAGLDTVSGILSAISASFILSNADADTGT 267

Qy       241 KVAAGFELSNQVIGNVTKAISSYVLAQRVAAGLSTTGAVAALITSSIMLAISPLAFMNAA 300
             | ||| ||:  :|:|||  |  ||  |::|||  |  ||||:  |   |  ||  |   :  ||||||:|:: |
Db       268 KAAAGVELTTKVLGNVGKGISQYIIAQRAAQGLSTSAAAAGLIASVVTLAISPLSFLSIA 327

Qy       301 DKFNHANALDEFAKQFRKFGYDGDHLLAEYQRGVGTIEASLTTISTALGAVSAGVSAAAV 360
             |||   ||  ::|::::|:|  |||||  |||  :  :   |  |:||||||||  |  :||:|:||||
Db       328 DKFKRANKIEEYSQRFKKLGYDGDSLLAAFHKETGAIDASLTTISTVLASVSSGISAAAT 387

Qy       361 GSAVGTPIALLVAGVTGLISGILEASKQAMFESVANRLQGKILEWEKQNGGQNYFDKGYD 420
             | ||  |:: ||   |||:||||||||||||||  ||:::    |  ||||::  |:|||: |||
Db       388 TSLVGAPVSALVGAVTGIISGILEASKQAMFEHVASKMADVIAEWEKKH-GKNYFENGYD 446

Qy       421 SRYAAYLANNLKFLSELNKELEAERVIAITQQRWDNNIGELAGITKLGERIKSGKAYADA 480
             :|:||:|  :|  :  ||:  |||    |  |  | |||   |||   |||||:|:  |::  |||:|  |
Db       447 ARHAAFLEDNFEILSQYNKEYSVERSVLITQQHWDTLIGELAGVTRNGDKTLSGKSYIDY 506

Qy       481 FEDGKKVEAG----SNITLDAKTGIIDISNSNGKKTQALHFTSPLLTAGTESRERLTNGK 536
             :|:|||::|                          |  |||:|        |:     |  :||||  |  |  |||  :||
Db       507 YEEGKRLEKEPDEFQKQVFDPLKGNIDLSVI--KSSTLLKFITPLLTPGKEIRERRQSGK 564

Qy       537 YSYINKLKFGRVKNWQVTD-GEASSKLDFSKVIQRVAETEGT-DEIGLIVNAKAGNDDIF 594
             | || :|      |   | |      :     |     |:| :||    :             || :  :     |:| :|
Db       565 YEYITELLVKGVDKWTVKGVQDKGSVYDYSNLIQHASVGNNQYREIRIESHLGDGDDKVF 624

Qy       595 VGQGKMNIDGGDGHDRVFYSK-DGGFGNITVDGTSATEAGSYTVNRKVARGD--IYHEVV 651
             :    |    ||    |   |||   |:|  |  |  |   |:   :|:|||  ||||:|||  |  |  ||   :   |||
Db       625 LSAGSANIYAGKGHDVVYYDKTDTGY--LTIDGTKATEAGNYTVTR-VLGGDVKVLQEVV 681

Qy       652 KRQETKVGKRTETIQYRDYELRKV-GYGYQSTDNLKSVEEVIGSQFNDVFKGSKFNDIFH 710
             |  ||    |||||    |||  ||       :  |         ||||  ||||:||:   |  ||||  ||||
Db       682 KEQEVSVGKRTEKTQYRSYEFTHINGTDLTETDNLYSVEELIGTNRADKFFGSKFTDIFH 741

Qy       711 SGEGDDLLDGGAGDDRLF---------------------------------------- 728
              :|||  ::|   |:|||:
Db       742 GADGDDHIEGNDGNDRLYGDKGNDTLRGGNGDDQLYGGDGNDKLTGGVGNNYLNGGDGDD 801

Qy       729 ---------------GGKGNDRLSGDEGDDLLDGGSGDDVLNGGAGNDVYIFRKGDGNDT 773
                            ||||||:|  |  ||  ||||||  |:|:|  || |||:|  :   |  |:
Db       802 ELQVQGNSLAKNVLSGGKGNDKLYGSEGADLLDGGEGNDLLKGGYGNDIYRYLSGYGHHI 861

Qy       774 LYDGTG-NDKLAFADANISDIMIERTKEGIIVKRND-------HSGSINIPRWYITSNLQ 825
             :  |   |  :|||:  ||  :   |:   :|      :|: :  :            |      |       |:      :
Db       862 IDDDGGKDDKLSLADIDFRDVAFKREGNDLIMYKAEGNVLSIGHKNGITFRNWF-----E 916

Qy       826 NYQSNKTDHKIEQLIGKDGSYITSDQIDKIL---QDKKDGTVITSQELKKLADENK---- 878
                :  ::|:|||:  |||   ||  |  :  |      |              :     :          || :
Db       917 KESGDISNHQIEQIFDKDGRVITPDSLKKAFEYQQSNNQANYVYGEYASTYADLDNLNPL 976

Qy       879 ----SQKLSAS---DI-----ASSLNKLVGSMALFGTA-NSVSSNA 911
                 |: :||:     |:        |:|| :|  |: :  |     ||:: |
Db       977 INEISKIISAAGNFDVKEERSAASLLQLSGNASDFSYGRNSITLTA 1022
```

RESULT 27
Q43892
ID   Q43892   PRELIMINARY;   PRT; 1055 AA.
AC   Q43892;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Leukotoxin.
GN   LKTA.
OS   Actinobacillus actinomycetemcomitans (Haemophilus
OS   actinomycetemcomitans).
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;

```
OC   Pasteurellaceae; Actinobacillus.
OX   NCBI_TaxID=714;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=JP2;
RA   Kolodrubetz J., Dailey T., Kraig E.;
RT   "Nucleotide sequence of the leukotoxin gene from actino-bacillus
RT   Actinomycetemcomitans: Homology to the A-hemolysin/leukotoxin gene
RT   family.";
RL   Infect. Immun. 58:920-929(1991).
DR   EMBL; X16829; CAA34731.1; -.
DR   InterPro; IPR001343; Hemlysn_Ca_bind.
DR   InterPro; IPR003995; RtxA.
DR   Pfam; PF00353; hemolysinCabind; 6.
DR   Pfam; PF02382; RTX; 1.
DR   PRINTS; PR00313; CABNDNGRPT.
DR   PRINTS; PR01488; RTXTOXINA.
DR   PROSITE; PS00330; HEMOLYSIN_CALCIUM; 1.
SQ   SEQUENCE   1055 AA;  113853 MW;  5331C396FA76669E CRC64;

  Query Match             39.1%;  Score 1816;  DB 2;  Length 1055;
  Best Local Similarity   40.2%;  Pred. No. 4e-73;
  Matches  413;  Conservative  185;  Mismatches  299;  Indels  130;  Gaps   23;

Qy          3 NINVIKSNIQAGLNSTKSGLKNLYLAIPKDYDPQKGGTLNDFIKAADELGIARLAEEPNH 62
Db         27 NIDAAKEQLQKALDKLGKTGKKLTLYIPKNY--KKGNGLTALIKAAQKLGIEVYHEGKDG 84

Qy         63 TETAKKSVDTVNQFLSLTQTGIAISATKLEKFLQ--KHSTNKLAKGLDSVENIDRKLGKA 120
Db         85 PALTNGILNTGKKLLGLTERGLTLFAPELDKWIQGNKHLSNSVG----STGNLTKAIDKV 140

Qy        121 SNVLSTLSSFLGTALAGIELDSLIK---KGDAAPDA-LAKASIDLINEIIGNLSQSTQTI 176
Db        141 QSVLGTLQAFLNTAFSGMDLDALIKARQNGKNVTDVQLAKASLNLINELIGTISSITNNV 200

Qy        177 EAFSSQLAKLGSTISQAKGFSNIGNKLQNL----NFSKTNLGLEIITGLLSGISAGFALA 232
Db        201 DTFSKQLNKLGEALGQVKHFGSFGDKLKNLPKLGNLGK---GLGALSGVLSAISAALLLA 257

Qy        233 DKNASTGKKVAAGFELSNQVIGNVTKAISSYVLAQRVAAGLSTTGAVAALITSSIMLAIS 292
Db        258 NKDADTATKAAAAAELTNKVLGNIGKAITQYLIAQRAAAGLSTTGPVAGLIASVVSLAIS 317

Qy        293 PLAFMNAADKFNHANALDEFAKQFRKFGYDGDHLLAEYQRGVGTIEASLTTISTALGAVS 352
Db        318 PLSFLGIAKQFDRARMLEEYSKRFKKFGYNGDSLLGQFYKNTGIADAAITTINTVLSAIA 377

Qy        353 AGVSAAAVGSAVGTPIALLVAGVTGLISGILEASKQAMFESVANRLQGKILEWEKQNGGQ 412
Db        378 AGVGAASAGSLVGAPIGLLVSAITSLISGILDASKQAVFEHIANQLADKIKAWENKY-GK 436

Qy        413 NYFDKGYDSRYAAYLANNLKFLSELNKELEAERVIAITQQRWDNNIGELAGITKLGERIK 472
Db        437 NYFENGYDARHSAFLEDSLKLFNELREKYKTENILSITQQGWDQRIGELAGITRNGDRIQ 496

Qy        473 SGKAYADAFEDGKKVEAGSN----ITLDAKTGIIDISNSNGKKTQALHFTSPLLTAGTES 528
Db        497 SGKAYVDYLKKGEELAKHSDKFTKQILDPIKGNIDLSGIKGSTT--LTFLNPLLTAGKEE 554

Qy        529 RERLTNGKYSYINKLKFGRVKNWQVTDGEASSKL-DFSKVIQR-VAETEGTDEIGLIVNA 586
Db        555 RKTRQSGKYEFITELKVKGRTDWKVKGVPNSNGVYDFSNLIQHAVTRDNKVLEARLIANL 614

Qy        587 KAGNDDIFVGQGKMNIDGGDGHDRVFYSKDGGFGNITVDGTSATEAGSYTVNRKVARGDI 646
Db        615 GAKDDYVFVGSGSTIVNAGDGYDVVDYSK-GRTGALTIDGRNATKAGQYKVERDLSGTQV 673

Qy        647 YHEVVKRQETKVGKRTETIQYRDYEL--RKVGYGYQSTDNLKSVEEVIGSQFNDVFKGSK 704
Db        674 LQETVSKQETKRGKVTDLLEYRNYKLDYYYTNKGFKAHDELNSVEEIIGSTLRDKFYGSK 733

Qy        705 FNDIFHSGEGDDL--------------------------LDGGAGDDRLF--------- 728
Db        734 FNDVFHGHDGDDLIYGYDGDDRLYGDNGNDEIHGGQGNDKLYGGAGNDRLFGEYGNNYLD 793

Qy        729 ------------------GGKGNDRLSGDEGDDLLDGGSGDDVLNGGAGNDVYIFRKGDG 770
Db        794 GGEGDDHLEGGNGSDILRGGSGNDKLFGNQGDDLLDGGEGDDQLAGGEGNDIYVYRKEYG 853

Qy        771 NDTLYDGTGN-DKLAFADANISDIMIERTKEGIIVKRNDHSGSINIPRWYITSN----LQ 825
Db        854 HHTITEHSGDKDKLSLANINLKDVSFERNGNDLLLKTNNRT-AVTFKGWFSKPNSSAGLD 912

Qy        826 NYQ-------------------------SNKTDHKIEQLIGKDGSYITSDQIDKILQDK 859
Db        913 EYQRKLLEYAPEKDRARLKRQFELQRGKVDKSLNNKVEEIIGKDGERITSQDIDNLFDKS 972

Qy        860 KDGTVITSQELKKL-ADENKSQKLSASDIASS-----------LNKLVGSMALFGTA-NS 906
Db        973 GNKKTISPQELAGLIKNKGKSSSLMSSSRSSSMLTQKSGLSNDISRIISATSGFGSSGKA 1032

Qy        907 VSSNALQ 913
Db       1033 LSASPLQ 1039


RESULT 28
O85101
ID   O85101      PRELIMINARY;      PRT;   998 AA.
AC   O85101;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Hemolysin.
GN   EHXA.
OS   Escherichia coli.
OG   Plasmid EHEC-hemolysin plasmid.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Escherichia.
OX   NCBI_TaxID=562;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=enterohemorrhagic EC920006;
RX   MEDLINE=98261495; PubMed=9596716;
RA   Boerlin P., Chen S., Colbourne J.K., Johnson R., De Grandis S.,
RA   Gyles C.;
RT   "Evolution of enterohemorrhagic escherichia coli hemolysin plasmids
RT   and the locus for enterocyte effacement in Shiga toxin-producing E.
RT   coli.";
RL   Infect. Immun. 66:2553-2561(1998).
DR   EMBL; AF043471; AAC24352.1; -.
DR   InterPro; IPR001343; Hemlysn_Ca_bind.
DR   InterPro; IPR003995; RtxA.
DR   Pfam; PF00353; hemolysinCabind; 6.
DR   Pfam; PF02382; RTX; 1.
DR   PRINTS; PR00313; CABNDNGRPT.
DR   PRINTS; PR01488; RTXTOXINA.
KW   Plasmid.
SQ   SEQUENCE   998 AA;  107095 MW;  E5A566B239DCBE55 CRC64;

  Query Match             37.9%;  Score 1762;  DB 2;  Length 998;
  Best Local Similarity   40.4%;  Pred. No. 9.7e-71;
  Matches  396;  Conservative  177;  Mismatches  302;  Indels  104;  Gaps   24;

Qy          8 KSNIQAGLNSTKSGLKNLYLAIPKDYDPQKGGTLNDFIKAADELGIARLAEEPNHTETAK 67
Db         18 KSAFNTASSSVRSAGKKLILLIPDNYEAQ-GVGINELVKAADELGIEIHRTERDDTAIAN 76

Qy         68 KSVDTVNQFLSLTQTGIAISATKLEKFLQKHSTNKLAKGL-DSVENIDRKLGKASNVLST 126
Db         77 QFFGTAEKVVGLTERGVAIFAPQLDKLLQKY--QKVGSKIGGTAENVGNNLGKAGTVLSA 134

Qy        127 LSSFLGTALAGIELDSLIKK----GDAAPDALAKASIDLINEIIGNLSQSTQTIEAFSSQ 182
Db        135 LQNFTGIALSGMALDELLRKQREGEDISQNDIAKSSIELINQLVDTVSSINSTVDSFSEQ 194
```

```
Qy    183 LAKLGSTISQAKGFSNIGNKLQNL-NFSKTNLGLEIITGLLSGISAGFALADKNASTGKK 241
Db    195 LNQLGSFLSSKPRLSSVGGKLQNLPDLGPLGDGLDVVSGILSAVSASFILGNSDAHTGTK 254

Qy    242 VAAGFELSNQVIGNVTKAISSYVLAQRVAAGLSTTGAVAALITSSIMLAISPLAFMNAAD 301
Db    255 AAAGIELTTQVLGNVGKAVSQYILAQRMAQGLSTTAASAGLITSAVMLAISPLSFLAIAD 314

Qy    302 KFNHANALDEFAKQFRKFGYDGDHLLAEYQRGVGTIEASLTTISTALGAVSAGVSAAAVG 361
Db    315 KFERAKQLESYSERFKKFNYEGDALLAAFHKESGAIDAALTTINTVLSSVSAGVSAASSA 374

Qy    362 SAVGTPIALLVAGVTGLISGILEASKQAMFESVANRLQGKILEWEKQNGGQNYFDKGYDS 421
Db    375 SLIGAPISMLVSALTGTISGILEASKQAMFEHVADKFAARINEWEKEH-GKNYFENGYDA 433

Qy    422 RYAAYLANNLKFLSELNKELEAERVIAITQQRWDNNIGELAGITKLGERIKSGKAYADAF 481
Db    434 RHAAFLEDSLSLLADFSRQHAVERAVAITQQHWDEKIGELAGITRNADRSQSGKAYINYL 493

Qy    482 EDGKKVEAG----SNITLDAKTGIIDISNSNGKKTQALHFTSPLLTAGTESRERLTNGKY 537
Db    494 ENGGLLEAQPKEFTQQVFDPQKGTIDL--STGNVSSVLTFVTPTFTPGEEVRERKQSGKY 551

Qy    538 SYINKLKFGRVKNWQVTDGEASSK--LDFSKVIQRVAETEGTDEIGLIVNAKAGNDDI-F 594
Db    552 EYMTSLIVNGKDTWSV-KGIKNHKGVYDYSNLIQFV-EKDNKHYQARIISELGDKDDIVY 609

Qy    595 VGQGKMNIDGGDGHDRVFYSKDGGFGNITVDGTSATEAGSYTVNRKVARGD--IYHEVVK 652
Db    610 SGAGSSEVFAGEGHDTVSYNKT-DVGKLTIDATGASKPGEYIVSKNM-YGDVKVLQEVVK 667

Qy    653 RQETKVGKRTETIQYRDYELRKVGYGYQSTDNLKSVEEVIGSQFNDVFKGSKFNDIFHSG 712
Db    668 EQEVSVGKRTEKIQYRDFEFRTGGIPYDVIDNLHSVEELIGGKHDDEFKGGKFNDIFHGA 727

Qy    713 -----------------EGDDLLDGGAGDDRLFGGKGNDRLSGDEGD----------- 742
Db    728 DGNDYIEGNYGNDRLYGDDGDDYISGGQGDDQLFGGSGNDKLSGGDGNNYLTGGSGNDEL 787

Qy    743 -----------------------DLLDGGSGDDVLNGGAGNDVYIFRKGDGNDTLYD-G 777
Db    788 QAHGAYNILSGGTGDDKLYGGGGIDLLDGGEGNDYLNGGFGNDIYVYRQNYGHHTIADEG 847

Qy    778 TGNDKLAFADANISDIMIERTKEGIIVKR----------NDHSGSINIPRWYITSNLQN 826
Db    848 GKGDRLHLSDISFDDIAFKKVGNDLIMNKAINGALSFNESNDVNG-ITFKNWFAKD---- 902

Qy    827 YQSNKTDHKIEQLIGKDGSYITSDQIDKILQDKKDGTVITSQELKKLADENKSQKLSASD 886
Db    903 -ASGEDNHLVEVITDKDGREIKADKISH-NNNEQSGYIKAS----NIASE--KNMVNITS 954

Qy    887 IASSLNKLVGSMALFGTAN 905
Db    955 VANDINKIISSVSGFDSGD 973
```

RESULT 29
Q9LC58
ID   Q9LC58   PRELIMINARY;   PRT;   998 AA.
AC   Q9LC58;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Hemolysin A.
GN   EHEC-HLYA.
OS   Escherichia coli.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Escherichia.
OX   NCBI_TaxID=562;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Taneike I., Wakisaka-Saito N., Harada Y., Zhang H.-M., Yamamoto T.;
RT   "The enterohemorrhagic Escherichia coli (EHEC)-hemolysin genes of a
RT   Shiga toxin 1 (Stx1)- and Stx2-producing, serotype O128 Escherichia
RT   coli strain with a greatest hemolytic activity.";
RL   Acta Med. Biol. (Niigata) 0:0-0(2000).
DR   EMBL; AB032930; BAA93708.1; -.
DR   InterPro; IPR001343; Hemlysn_Ca_bind.
DR   InterPro; IPR003995; RtxA.
DR   Pfam; PF00353; hemolysinCabind; 6.
DR   Pfam; PF02382; RTX; 1.
DR   PRINTS; PR00313; CABNDNGRPT.
DR   PRINTS; PR01488; RTXTOXINA.
DR   PROSITE; PS00330; HEMOLYSIN_CALCIUM; 1.
SQ   SEQUENCE   998 AA; 107196 MW; CD7A88E9BD862DB6 CRC64;

Query Match 37.9%; Score 1760; DB 2; Length 998;
Best Local Similarity 40.4%; Pred. No. 1.2e-70;
Matches 396; Conservative 174; Mismatches 305; Indels 104; Gaps 24;

```
Qy      8 KSNIQAGLNSTKSGLKNLYLAIPKDYDPQKGGTLNDFIKAADELGIARLAEEPNHTETAK 67
Db     18 KSAFNTASSSVRSAGKKLILLIPDNYEAQ-GVGINELVKAADDLGIEIHRTERDDTAIAN 76

Qy     68 KSVDTVNQFLSLTQTGIAISATKLEKFLQKHSTNKLAKGL-DSVENIDRKLGKASNVLST 126
Db     77 QFFGTAEKVVGLTERGVAIFAPQLDKLLQKY--QKVGSKIGGTAENVGNNLGKAGTVLSA 134

Qy    127 LSSFLGTALAGIELDSLIKK----GDAAPDALAKASIDLINEIIGNLSQSTQTIEAFSSQ 182
Db    135 LQNFTGIALSGMALDELLRKQREGEDISQNDIAKSSIELINQLVDTVSSINSTVDSFSEQ 194

Qy    183 LAKLGSTISQAKGFSNIGNKLQNL-NFSKTNLGLEIITGLLSGISAGFALADKNASTGKK 241
Db    195 LNQLGSFLSSKPRLSSVGGKLQNLPNLGPLGDGLDVVSGILSAVSASFILGNSDAHTGTK 254

Qy    242 VAAGFELSNQVIGNVTKAISSYVLAQRVAAGLSTTGAVAALITSSIMLAISPLAFMNAAD 301
Db    255 AAAGIELTTQVLGNVGKAVSQYILAQRMAQGLSTTAASAGLITSAVMLAISPLSFLAIAD 314

Qy    302 KFNHANALDEFAKQFRKFGYDGDHLLAEYQRGVGTIEASLTTISTALGAVSAGVSAAAVG 361
Db    315 KFERAKQLESYSERFKKLNYEGDALLAAFHKETGAIDAALTTINTVLSSVSAGVSAASSA 374

Qy    362 SAVGTPIALLVAGVTGLISGILEASKQAMFESVANRLQGKILEWEKQNGGQNYFDKGYDS 421
Db    375 SLIGAPISMLVSALTGTISGILEASKQAMFEHVADKFAARINEWEKEH-GKNYFENGYDA 433

Qy    422 RYAAYLANNLKFLSELNKELEAERVIAITQQRWDNNIGELAGITKLGERIKSGKAYADAF 481
Db    434 RHAAFLEDSLSLLADFSRQHAVERAVAITQQHWDEKIGELAGITRNADRSQSGKAYINYL 493

Qy    482 EDGKKVEAG----SNITLDAKTGIIDISNSNGKKTQALHFTSPLLTAGTESRERLTNGKY 537
Db    494 ENGGLLEAQPKEFTQQVFDPQKGTIDL--STGNVSSVLTFVTPTFTPGEEVRERKQSGKY 551

Qy    538 SYINKLKFGRVKNWQVTDGEASSK--LDFSKVIQRVAETEGTDEIGLIVNAKAGNDDI-F 594
Db    552 EYMTSLIVNGKDTWSV-KGIKNHKGVYDYSNLIQFV-EKDNKHYQARIISELGDKDDIVY 609

Qy    595 VGQGKMNIDGGDGHDRVFYSKDGGFGNITVDGTSATEAGSYTVNRKVARGD--IYHEVVK 652
Db    610 SGAGSSEVFAGEGHDTVSYNKT-DVGKLTIDATGASKPGEYIVSKNM-YGDVKVLQEVVK 667

Qy    653 RQETKVGKRTETIQYRDYELRKVGYGYQSTDNLKSVEEVIGSQFNDVFKGSKFNDIFHSG 712
Db    668 EQEVSVGKRTEKIQYRDFEFRTGGIPYDVIDNLHSVEELIGGKHDDEFKGGKFNDIFHGA 727

Qy    713 EG-----------------DDLLDGGAGDDRLFGGKGNDRLSGDEGD----------- 742
Db    728 DGNDYIEGNYGNDRLYGDDRDDYISGGQGDDQLFGGSGNDKLSGGDGNNYLTGGSGNDEL 787

Qy    743 -----------------------DLLDGGSGDDVLNGGAGNDVYIFRKGDGNDTLYD-G 777
```

```
Db        788 QAHGAYNILSGGTGDDKLYGGGGIDLLDGGEGNDYLNGGFGNDIYVYRQNYGHHTIADEG 847

Qy        778 TGNDKLAFADANISDIMIERTKEGIIVKR-----------NDHSGSINIPRWYITSNLQN 826

Db        848 GKGDRLHLSDISFDDIAFKRVGNDLIMNKAINGVLSFNESNDVNG-ITFKNWFAKD---- 902

Qy        827 YQSNKTDHKIEQLIGKDGSYITSDQIDKILQDKKDGTVITSQELKKLADENKSQKLSASD 886

Db        903 -ASGADNHLVEVITDKDGREIKADKIPH-NNNERSGYIKAS----NIASE--KNMVNITS 954

Qy        887 IASSLNKLVGSMALFGTAN 905

Db        955 VADDINKIISSVSGFDSGD 973

RESULT 30
P71223
ID   P71223      PRELIMINARY;      PRT;   998 AA.
AC   P71223;
DT   01-FEB-1997 (TrEMBLrel. 02, Created)
DT   01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   EHEC-hemolysin.
GN   EHEC-HLYA.
OS   Escherichia coli.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Escherichia.
OX   NCBI_TaxID=562;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=78-92;
RX   MEDLINE=97034824; PubMed=8880480;
RA   Schmidt H., Karch H.;
RT   "Enterohemolytic phenotypes and genoty[es pof Shiga Toxin-producing
RT   Esherichia coli 0111 strain from patients with diarrhea and hemolytic-
RT   uremic syndrome.";
RL   J. Clin. Microbiol. 34:2364-2367(1996).
DR   EMBL; X94129; CAA63849.1; -.
DR   InterPro; IPR001343; Hemlysn_Ca_bind.
DR   InterPro; IPR003995; RtxA.
DR   Pfam; PF00353; hemolysinCabind; 6.
DR   Pfam; PF02382; RTX; 1.
DR   PRINTS; PR00313; CABNDNGRPT.
DR   PRINTS; PR01488; RTXTOXINA.
SQ   SEQUENCE   998 AA;  107255 MW;  DA3EF078C7E4131E CRC64;

  Query Match             37.8%;  Score 1756;  DB 2;  Length 998;
  Best Local Similarity   40.1%;  Pred. No. 1.8e-70;
  Matches  393;  Conservative  176;  Mismatches  306;  Indels  104;  Gaps   21;

Qy          8 KSNIQAGLNSTKSGLKNLYLAIPKDYDPQKGGTLNDFIKAADELGIARLAEEPNHTETAK 67

Db         18 KSAFNTASSSVRSAGKKLILLIPDNYEAQ-GVGINELVKAADELGIEIHRTERDDTAIAN 76

Qy         68 KSVDTVNQFLSLTQTGIAISATKLEKFLQKHST--NKLAKGLDSVENIDRKLGKASNVLS 125

Db         77 QFFGAAEKVVGLTERGVAIFAPQLDKLLQKYQKVGSKIGR---TAENVGNNLGKAGTVLS 133

Qy        126 TLSSFLGTALAGIELDSLIKK----GDAAPDALAKASIDLINEIIGNLSQSTQTIEAFSS 181

Db        134 ALQNFTGIALSGMALDELLRKQREGEDISQNDIAKSSIELINQLVDTVSSINSTVDSFSE 193

Qy        182 QLAKLGSTISQAKGFSNIGNKLQNL-NFSKTNLGLEIITGLLSGISAGFALADKNASTGK 240

Db        194 QLNQLGSFLSSKPRLSSVGGKLQNLPDLGSLGDGLDVVSGILSAVSASFILGNSDAHTGT 253

Qy        241 KVAAGFELSNQVIGNVTKAISSYVLAQRVAAGLSTTGAVAALITSSIMLAISPLAFMNAA 300

Db        254 KAAAGIELTTQVLGNVGKAVSQYILAQRMAQGLSTTAASAGLITSAVMLAISPLSFLAAA 313

Qy        301 DKFNHANALDEFAKQFRKFGYDGDHLLAEYQRGVGTIEASLTTISTALGAVSAGVSAAAV 360

Db        314 DKFERAKQLESYSERFKKLNYEGDALLAAFHKETGAIDAALTTINTVLSSVSAGVSAASS 373

Qy        361 GSAVGTPIALLVAGVTGLISGILEASKQAMFESVANRLQGKILEWEKQNGGQNYFDKGYD 420

Db        374 ASLIGAPISMLVSALTGTISGILEASKQAMFEHVAEKFAARINEWEKEH-GKNYFENGYD 432

Qy        421 SRYAAYLANNLKFLSELNKELEAERVIAITQQRWDNNIGELAGITKLGERIKSGKAYADA 480

Db        433 ARHAAFLEDSLSLLADFSRQHAVERAVAITQQHWDEKIGELAGITRNADRSQSGKAYINY 492

Qy        481 FEDGKKVEAG----SNITLDAKTGIIDISNSNGKKTQALHFTSPLLTAGTESRERLTNGK 536

Db        493 LENGGLLEAQPKEFTQQVFDPQKGTIDL--STGNVSSVLTFITPTFTPGEEVRERKQSGK 550

Qy        537 YSYINKLKFGRVKNWQVTDGEASSK--LDFSKVIQRVAETEGTDEIGLIVNAKAGNDDIF 594

Db        551 YEYMTFLIVNGKDTWSV-KGIKNHKGVYDYSKLIQFVEKDTKHYQARMISELGDKDDVVY 609

Qy        595 VGQGKMNIDGGDGHDRVFYSKDGGFGNITVDGTSATEAGSYTVNRKVARGD--IYHEVVK 652

Db        610 SGAGSSEVFAGEGHDPVSYNKT-DVGKLTIDATGASKPGEYIVPKNM-YGDVEVLQEVVK 667

Qy        653 RQETKVGKRTETIQYRDYELRKVGYGYQSTDNLKSVEEVIGSQFNDVFKGSKFNDIFHSG 712

Db        668 EQEVSVGKRTEKIQYRDFEFRTGGIPYDVIDNLHSVEELIGGKHDDEFKGGKFNDIFHGA 727

Qy        713 -----------------EGDDLLDGGAGDDRLFGGKGNDRLSGDEGD------------ 742

Db        728 DGNDYIEGNYGNDRLYGDDGDDYISGGQGDDQLFGGSGNDKLSGGDGNNYLTGGSGNDEL 787

Qy        743 ------------------------DLLDGGSGDDVLNGGAGNDVYIFRKGDGNDTLYD-G 777

Db        788 QAHGAYNILSGGTGDDKLYGGGGIDLLDGGEGNDYLNGGFGNDIYVYRQNYGHHTIADEG 847

Qy        778 TGNDKLAFADANISDIMIERTKEGIIVKR-----------NDHSGSINIPRWYITSNLQN 826

Db        848 GKGDRLHLSDISFDDIAFKRVGNDLIMNKAINGVLSFNESNDVNG-ITFKNWFAKD---- 902

Qy        827 YQSNKTDHKIEQLIGKDGSYITSDQIDKILQDKKDGTVITSQELKKLADENKSQKLSASD 886

Db        903 -ASGADNHLVEVITDKDGREI---KVDKIPHNNNE----RSGYIKASNIASEKSMVNITS 954

Qy        887 IASSLNKLVGSMALFGTAN 905

Db        955 VANDINKIISSVSGFDSGD 973

RESULT 31
Q46716
ID   Q46716      PRELIMINARY;      PRT;   998 AA.
AC   Q46716;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Hemolysin A (HLYA).
GN   HLYA OR EHEC-HLYA.
OS   Escherichia coli O157:H7.
OG   Plasmid pO157.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Escherichia.
OX   NCBI_TaxID=83334;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98290540; PubMed=9628576;
RA   Makino K., Ishii K., Yasunaga T., Hattori M., Yokoyama K.,
RA   Yatsudo H.C., Kubota Y., Yamaichi Y., Iida T., Yamamoto K., Honda T.,
RA   Han C., Ohtsubo A., Kasamatsu M., Hayashi T., Kuhara S.,
RA   Shinagawa H.;
RT   "Complete nucleotide sequences of 93-kb and 3.3-kb plasmids of an
RT   enterohemorrhagic Escherichia coli O157:H7 derived from Sakai
RT   outbreak.";
RL   DNA Res. 5:1-9(1998).
RN   [2]
RP   SEQUENCE OF 241-998 FROM N.A.
```

```
RC   STRAIN=EHEC;
RA   Hall R.H., Xu J., Walderhaug M.O.;
RL   Submitted (JUL-1994) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=EDL933;
RX   MEDLINE=98391744; PubMed=9722640;
RA   Burland V., Shao Y., Perna N.T., Plunkett G., Sofia H.J.,
RA   Blattner F.R.;
RT   "The complete DNA sequence and analysis of the large virulence plasmid
RT   of Escherichia coli O157:H7.";
RL   Nucleic Acids Res. 26:4196-4204(1998).
DR   EMBL; AB011549; BAA31774.1; -.
DR   EMBL; U12572; AAA20544.1; -.
DR   EMBL; AF074613; AAC70116.1; -.
DR   InterPro; IPR001343; Hemlysn_Ca_bind.
DR   InterPro; IPR003995; RtxA.
DR   Pfam; PF00353; hemolysinCabind; 6.
DR   Pfam; PF02382; RTX; 1.
DR   PRINTS; PR00313; CABNDNGRPT.
DR   PRINTS; PR01488; RTXTOXINA.
KW   Plasmid.
SQ   SEQUENCE   998 AA;  107047 MW;  BB3C0D36FCBB9EBD CRC64;

  Query Match             37.6%;  Score 1748;  DB 2;  Length 998;
  Best Local Similarity   40.0%;  Pred. No. 4.1e-70;
  Matches  391;  Conservative  177;  Mismatches  308;  Indels  102;  Gaps   21;

Qy          8 KSNIQAGLNSTKSGLKNLYLAIPKDYDPQKGGTLNDFIKAADELGIARLAEEPNHTETAK 67
Db         18 KSAFNTASSSVRSAGKKLILLIPDNYEAQ-GVGINELVKAADELGIEIHRTERDDTAIAN 76

Qy         68 KSVDTVNQFLSLTQTGIAISATKLEKFLQKHSTNKLAKGL-DSVENIDRKLGKASNVLST 126
Db         77 QFFGAAEKVVGLTERGVAIFAPQLDKLLQKY--QKVGSKIGGTAENVGNNLGKAGTVLSA 134

Qy        127 LSSFLGTALAGIELDSLIKK----GDAAPDALAKASIDLINEIIGNLSQSTQTIEAFSSQ 182
Db        135 LQNFTGIALSGMALDELLRKQREGEDISQNDIAKSSIELINQLVDTVSSINSTVDSFSEQ 194

Qy        183 LAKLGSTISQAKGFSNIGNKLQNL-NFSKTNLGLEIITGLLSGISAGFALADKNASTGKK 241
Db        195 LNQLGSFLSSKPRLSSVGGKLQNLPDLGPLGDGLDVVSGILSAVSASFILGNSDAHTGTK 254

Qy        242 VAAGFELSNQVIGNVTKAISSYVLAQRVAAGLSTTGAVAALITSSIMLAISPLAFMNAAD 301
Db        255 AAAGIELTTQVLGNVGKAVSQYILAQRMAQGLSTTAASAGLITSAVMLAISPLSFLAAAD 314

Qy        302 KFNHANALDEFAKQFRKFGYDGDHLLAEYQRGVGTIEASLTTISTALGAVSAGVSAAAVG 361
Db        315 KFERAKQLESYSERFKKLNYEGDALLAAFHKETGAIDAALTTINTVLSSVSAGVSAASSA 374

Qy        362 SAVGTPIALLVAGVTGLISGILEASKQAMFESVANRLQGKILEWEKQNGGQNYFDKGYDS 421
Db        375 SLIGAPISMLVSALTGTISGILEASKQAMFEHVAEKFAARINEWEKEH-GKNYFENGYDA 433

Qy        422 RYAAYLANNLKFLSELNKELEAERVIAITQQRWDNNIGELAGITKLGERIKSGKAYADAF 481
Db        434 RHAAFLEDSLSLLADFSRQHAVERAVAITQQHWDEKIGELAGITRNADRSQSGKAYINYL 493

Qy        482 EDGKKVEAG----SNITLDAKTGIIDISNSNGKKTQALHFTSPLLTAGTESRERLTNGKY 537
Db        494 ENGGLLEAQPKEFTQQVFDPQKGTIDL--STGNVSSVLTFITPTFTPGEEVRERKQSGKY 551

Qy        538 SYINKLKFGRVKNWQVTDGEASSK--LDFSKVIQRVAETEGTDEIGLIVNAKAGNDDIFV 595
Db        552 EYMTSLIVNGKDTWSV-KGIKNHKGVYDYSKLIQFVEKNNKHYQARIISELGDKDDVVYS 610

Qy        596 GQGKMNIDGGDGHDRVFYSKDGGFGNITVDGTSATEAGSYTVNRKVARGD--IYHEVVKR 653
Db        611 GAGSSEVFAGEGYDTVSYNKT-DVGKLTIDATGASKPGEYIVSKNM-YGDVKVLQEVVKE 668

Qy        654 QETKVGKRTETIQYRDYELRKVGYGYQSTDNLKSVEEVIGSQFNDVFKGSKFNDIFHSG- 712
Db        669 QEVSVGKRTEKIQYRDFEFRTGGIPYDVIDNLHSVEELIGGKHDDEFKGGKFNDIFHGAD 728

Qy        713 -----------------EGDDLLDGGAGDDRLFGGKGNDRLSGDEGD------------- 742
Db        729 GNDYIEGNYGNDRLYGDDGDDYISGGQGDDQLFGGSGNDKLSGDGNNYLTGGSGNDELQ 788

Qy        743 ----------------------DLLDGGSGDDVLNGGAGNDVYIFRKGDGNDTLYD-GT 778
Db        789 AHGAYNILSGGTGDDKLYGGGGIDLLDGGEGNDYLNGGFGNDIYVYGQNYGHHTIADEGG 848

Qy        779 GNDKLAFADANISDIMIERTKEGIIVKR-----------NDHSGSINIPRWYITSNLQNY 827
Db        849 KGDRLHLSDISFDDIAFKRVGNDLIMNKAINGVLSFNESNDVNG-ITFKNWFAKD----- 902

Qy        828 QSNKTDHKIEQLIGKDGSYITSDQIDKILQDKKDGTVITSQELKKLADENKSQKLSASDI 887
Db        903 ASGADNHLVEVITDKDGREI---KVDKIPHNNNE----RSGYIKASNIASEKNMVNITSV 955

Qy        888 ASSLNKLVGSMALFGTAN 905
Db        956 ANDINKIISSVSGFDSGD 973

RESULT 32
Q93NP2
ID   Q93NP2     PRELIMINARY;     PRT;   1022 AA.
AC   Q93NP2;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   RTX toxin IA.
OS   Actinobacillus pleuropneumoniae (Haemophilus pleuropneumoniae).
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC   Pasteurellaceae; Actinobacillus.
OX   NCBI_TaxID=715;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Shin S.J., Park J.Y., Choi I.S., Shin N.R., Yoo H.S.;
RT   "Cloning and Sequencing of Apx IA from Actinobacillus
RT   pleuropneumoniae.";
RL   Submitted (MAR-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF363361; AAK50051.1; -.
DR   InterPro; IPR001343; Hemlysn_Ca_bind.
DR   InterPro; IPR003995; RtxA.
DR   Pfam; PF00353; hemolysinCabind; 6.
DR   Pfam; PF02382; RTX; 1.
DR   PRINTS; PR00313; CABNDNGRPT.
DR   PRINTS; PR01488; RTXTOXINA.
SQ   SEQUENCE   1022 AA;  110439 MW;  A68E486D335467E3 CRC64;

  Query Match             37.4%;  Score 1736;  DB 2;  Length 1022;
  Best Local Similarity   39.4%;  Pred. No. 1.4e-69;
  Matches  400;  Conservative  178;  Mismatches  317;  Indels  120;  Gaps   26;

Qy          8 KSNIQAGLNSTKSGL-------KNLYLAIPKDYDPQKGGTLNDFIKAADELGIARLAEEP 60
Db         21 KSAAKSGAGALKNGLGQVKQAGQKLILYIPKDYQASTGSSLNDLVKAAEALGIEVHRSEK 80

Qy         61 NHTETAKKSVDTVNQFLSLTQTGIAISATKLEKFLQKHSTNKLAKGL-DSVENIDRKLGK 119
Db         81 NGTALAKELFGTTEKLLGFSERGIALFAPQFDKLLNKN--QKLSKSLGGSSEALGQRLNK 138

Qy        120 ASNVLSTLSSFLGTALAGIELDSLIKK----GDAAPDALAKASIDLINEIIGNLSQSTQT 175
Db        139 TQTALSALQSFLGTAIAGMDLDSLLRRRRNGEDVSGSELAKAGVDLAAQLVDNIASATGT 198

Qy        176 IEAFSSQLAKLGSTISQAKGFSNIGNKLQNL-NFSKTNLGLEIITGLLSGISAGFALADK 234
Db        199 VDAFAEQLGKLGNALSNTR-LSGLASKLNNLPDLSLAGPGFDAVSGILSVVSASFILSNK 257

Qy        235 NASTGKKVAAGFELSNQVIGNVTKAISSYVLAQRVAAGLSTTGAVAALITSSIMLAISPL 294
```

```
Db        258 DADAGTKAAAGIEISTKILGNIGKAVSQYIIAQRVAAGLSTTAATRWFNRSVVALAISPL 317

Qy        295 AFMNAADKFNHANALDEFAKQFRKFGYDGDHLLAEYQRGVGTIEASLTTISTALGAVSAG 354
Db        318 SFLNVADKFERAKQLEQYSERFKKFGYEGDSLLASFYRETGAIEAALTTINSVLSAAPAG 377

Qy        355 VSAAAVGSAVGTPIALLVAGVTGLISGILEASKQAMFESVANRLQGKILEWEKQNGGQNY 414
Db        378 VGAAATGSLVGAPVAALVSAITGIISGILDASKQAIFERVATKLANKIDEWEKKH-GKNY 436

Qy        415 FDKGYDSRYAAYLANNLKFLSELNKELEAERVIAITQQRWDNNIGELAGITKLGERIKSG 474
Db        437 FENGYDARHSAFLEDTFELLSQYNKEYSVERVVAITQQRWDVNIGELAGITRKGADAKSG 496

Qy        475 KAYADAFEDGKKVEAG----SNITLDAKTGIIDISNSNGKKTQALHFTSPLLTAGTESRE 530
Db        497 KAYVDFFEEGKLLEKDPDRFDKKVFDPLEGKIDLSSIN--KTTLLKFITPAFTAGEEIRE 554

Qy        531 RLTNGKYSYINKLKFGRVKNWQVTDGEASSKL-DFSKVIQRVAETEGTDEIGLIVNAKAG 589
Db        555 RKQTGKYEYMTELFVKGKEKWVVTGVQSHNAIYDYTNLIQLAIDKKG-EKRQVTIESHLG 613

Qy        590 --NDDIFVGQGKMNIDGGDGHDRVFYSK-DGGFGNITVDGTSATEAGSYTVNRKV-ARGD 645
Db        614 EKNDRIYLSSGSSIEYAGNGHDVAYYDKTDTGY--LTFDGQSAQKAGEYIVTKELKADVK 671

Qy        646 IYHEVVKRQETKVGKRTETIQYRDYELR--KVGYGYQSTDNLKSVEEVIGSQFNDVFKGS 703
Db        672 VLKEVVKTQDISVGKRSEKLEYRDYELSPFELGNGIRAKDELHSVEEIIGSNRKDKFFGS 731

Qy        704 KFNDIFHSGEGD-----------------DLLDGGAGDDRLFGGKGNDRLSGDEGDDLL 745
Db        732 RFTDIFHGAKGDDEIYGNDGHDILYGDDGNDVIHGGDGNDHLVGGNGNDRLIGGKGNNFL 791

Qy        746 DGGSGDD----------VLNGGAGNDVY-------IFRKGDGNDTLYDGTGND------- 781
Db        792 NGGDGDDELQVFECQYNVLLGGAGNDILYGSDGTNLFDGGVGNDKIYGGLGKDIYRYSKE 851

Qy        782 ------------------KLAFADANI----SDIMIERTKEGIIVKRNDHSG-SINIPR 817
Db        852 YGRHIIIEKGGDDDTLLLSDLSFKDVGFIRIGDDLLVNKRIGGTLYYHEDYNGNALTIKD 911

Qy        818 WYITSNLQNYQSNKTDHKIEQLIGKDGSYITSDQIDKILQDKKDGTVITSQELKKLADEN 877
Db        912 WF--KEGKEGQNN----KVEKIVDKDGAYVLSQYLTELTAPGRGINYFNGLEEKLYYGEG 965

Qy        878 KSQ----KLSASDIASSLNKLVGS----------MALFGTANSVSSNALQPITQ 917
Db        966 YNALPQLRKDIEQIISSTGALTGEHGQVLVGAGGPLAYSNSPNSIPNAFSNYLTQ 1020
```

RESULT 33
Q47461
ID   Q47461   PRELIMINARY;   PRT;   998 AA.
AC   Q47461;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   EHEC-hlyA protein.
GN   EHEC-HLYA.
OS   Escherichia coli.
OG   Plasmid p0157.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Escherichia.
OX   NCBI_TaxID=562;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=EDL 933;
RX   MEDLINE=95172699; PubMed=7868227;
RA   Schmidt H., Beutin L., Karch H.;
RT   "Molecular analysis of the plasmid-encoded hemolysin of Escherichia
RT   coli 0157:H7 strain Edl933.";
RL   Infect. Immun. 63:1055-1061(1995).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=EDL 933;
RX   MEDLINE=97090409; PubMed=8936317;
RA   Schmidt H., Kernbach C., Karch H.;
RT   "Analysis of the EHEC hly operon and its location in the physical map
RT   of the large plasmid of enterohemorrhagic Esherichia coli 0157:H7.";
RL   Microbiology 142:907-914(1996).
DR   EMBL; X86087; CAA60042.1; -.
DR   InterPro; IPR001343; Hemlysn_Ca_bind.
DR   InterPro; IPR003995; RtxA.
DR   Pfam; PF00353; hemolysinCabind; 6.
DR   Pfam; PF02382; RTX; 1.
DR   PRINTS; PR00313; CABNDNGRPT.
DR   PRINTS; PR01488; RTXTOXINA.
KW   Plasmid.
SQ   SEQUENCE   998 AA; 107032 MW; 4D8BE108C309BF7E CRC64;

Query Match 37.3%; Score 1734; DB 2; Length 998;
Best Local Similarity 39.9%; Pred. No. 1.7e-69;
Matches 390; Conservative 174; Mismatches 312; Indels 102; Gaps 21;

```
Qy          8 KSNIQAGLNSTKSGLKNLYLAIPKDYDPQKGGTLNDFIKAADELGIARLAEEPNHTETAK 67
Db         18 KSAFNTASSSVRSAGKKLILLIPDNYEAQ-GVGINELVKAADELGIEIHRTERDDTAIAN 76

Qy         68 KSVDTVNQFLSLTQTGIAISATKLEKFLQKHSTNKLAKGL-DSVENIDRKLGKASNVLST 126
Db         77 QFFGAAEKVVGLTERGVAIFAPQLDKLLQKY--QKVGSKIGGTAENVGNNLGKAGTVLSA 134

Qy        127 LSSFLGTALAGIELDSLIKK----GDAAPDALAKASIDLINEIIGNLSQSTQTIEAFSSQ 182
Db        135 LQNFTGIALSGMALDELLRKQREGEDISQNDIAKSSIELINQLVDTVSSINSTVDSFSEQ 194

Qy        183 LAKLGSTISQAKGFSNIGNKLQNL-NFSKTNLGLEIITGLLSGISAGFALADKNASTGKK 241
Db        195 LNQLGSFLSSKPRLSSVGGKLQNLPDLGPLGDGLDVVSGILSAVSASFILGNSDAHTGTK 254

Qy        242 VAAGFELSNQVIGNVTKAISSYVLAQRVAAGLSTTGAVAALITSSIMLAISPLAFMNAAD 301
Db        255 AAAGIELTTQVLGNVGKAVSQYILAQRMAQGLSTTAASAGLITSAVMLAISPLSFLAAAD 314

Qy        302 KFNHANALDEFAKQFRKFGYDGDHLLAEYQRGVGTIEASLTTISTALGAVSAGVSAAAVG 361
Db        315 KFERAKQLESYSERFKKLNYEGDALLAGFHKETGAIDAGLTTINTVLSSVSAGVSAASSA 374

Qy        362 SAVGTPIALLVAGVTGLISGILEASKQAMFESVANRLQGKILEWEKQNGGQNYFDKGYDS 421
Db        375 SLIGAPISMLVSALTGTISGILEASKQAMFEHVAEKFAARINEWEKEH-GKNYFENGYDA 433

Qy        422 RYAAYLANNLKFLSELNKELEAERVIAITQQRWDNNIGELAGITKLGERIKSGKAYADAF 481
Db        434 RHAAFLEDSLSLLADFSRQHAVERAVAITHQHWDEKIGELAGITRNADRSQSGKAYINYL 493

Qy        482 EDGKKVEAG----SNITLDAKTGIIDISNSNGKKTQALHFTSPLLTAGTESRERLTNGKY 537
Db        494 ENGGLLEAQPKEFTQQVFDPQKGTIDL--STGNVSSVLTFITPTFTPGEEVRERKQSGKY 551

Qy        538 SYINKLKFGRVKNWQVTDGEASSK--LDFSKVIQRVAETEGTDEIGLIVNAKAGNDDIFV 595
Db        552 EYMTSLIVNGKDTWSV-KGIKNHKGVYDYSKLIQFVEKNTKHYQARIISELGDKDDVVYS 610

Qy        596 GQGKMNIDGGDGHDRVFYSKDGGFGNITVDGTSATEAGSYTVNRKVARGD--IYHEVVKR 653
Db        611 GAGSSEVFAGEGYDPVSYNKT-DVGKLTIDATGAPKPGEYIVPKNM-YGDVEVLQEVVKE 668

Qy        654 QETKVGKRTETIQYRDYELRKVGYGYQSTDNLKSVEEVIGSQFNDVFKGSKFNDIFHSG- 712
Db        669 QEVSVGKRTEKIQYRDFEFRTGGIPYDVIDNLHSVEELIGGKHDDEFKGGKFNDIFHGAD 728

Qy        713 ----------------EGDDLLDGGAGDDRLFGGKGNDRLSGDEGD------------- 742
Db        729 GNDYIEGNYGNDRLYGDDGDDYISGGQGDDQLFGGSGNDKLSGGDGNNYLTGGSGNDELQ 788
```

```
Qy        743 ----------------------DLLDGGSGDDVLNGGAGNDVYIFRKGDGNDTLYD-GT 778
                                    ||||| |:| |||| |||:|:: :  |: |: | |
Db        789 AHGAYNILSGGTGDDKLYGGGGIDLLDGGEGNDYLNGGFGNDIYVYGQNYGHHTIADEGG 848

Qy        779 GNDKLAFADANISDIMIERTKEGIIVKR-----------NDHSGSINIPRWYITSNLQNY 827
                 |:|  :| :  ||  :|     :|: :            || :| |    |:
Db        849 KGDRLHLSDISFDDIAFKRVGNDLIMNKAINGVLSFNESNDVNG-ITFKNWFAKD----- 902

Qy        828 QSNKTDHKIEQLIGKDGSYITSDQIDKILQDKKDGTVITSQELKKLADENKSQKLSASDI 887
                |    :| :| :  |||  |    ::|||  :   :      | :|     ::   :: : :
Db        903 ASGADNHLVEVITDKDGREI---KVDKIPHNNNE----RSGYIKASNIASEKNMVNITSV 955

Qy        888 ASSLNKLVGSMALFGTAN 905
              |: :||:: |:: | : :
Db        956 ANDINKIISSVSGFDSGD 973
```

```
RESULT 34
Q47262
ID   Q47262      PRELIMINARY;      PRT;   998 AA.
AC   Q47262;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Hemolysin.
GN   EHEC-HLYA.
OS   Escherichia coli.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Escherichia.
OX   NCBI_TaxID=562;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=95172699; PubMed=7868227;
RA   Schmidt H., Beutin L., Karch H.;
RT   "Molecular analysis of the plasmid-encoded hemolysin of Escherichia
RT   coli 0157:H7 strain Edl933.";
RL   Infect. Immun. 63:1055-1061(1995).
DR   EMBL; X79839; CAA56234.1; -.
DR   InterPro; IPR001343; Hemlysn_Ca_bind.
DR   InterPro; IPR003995; RtxA.
DR   Pfam; PF00353; hemolysinCabind; 6.
DR   Pfam; PF02382; RTX; 1.
DR   PRINTS; PR00313; CABNDNGRPT.
DR   PRINTS; PR01488; RTXTOXINA.
SQ   SEQUENCE   998 AA;  107058 MW;  0D3BE108C309B8B3 CRC64;

  Query Match             37.2%;  Score 1729;  DB 2;  Length 998;
  Best Local Similarity   39.8%;  Pred. No. 2.9e-69;
  Matches  389;  Conservative  174;  Mismatches  313;  Indels  102;  Gaps   21;

Qy          8 KSNIQAGLNSTKSGLKNLYLAIPKDYDPQKGGTLNDFIKAADELGIARLAEEPNHTETAK 67
               ||      :| :|  | | |  || :|: | |   :|: :|||||||      | : |  |
Db         18 KSAFNTASSSVRSAGKKLILLIPDNYEAQ-GVGINELVKAADELGIEIHRTERDDTAIAN 76

Qy         68 KSVDTVNQFLSLTQTGIAISATKLEKFLQKHSTNKLAKGL-DSVENIDRKLGKASNVLST 126
              :      : : ||: |:|| | :|:| |||:    |:    :   : ||:    ||||  |||
Db         77 QFFGAAEKVVGLTERGVAIFAPQLDKLLQKY--QKVGSKIGGTAENVGNNLGKAGTVLSA 134

Qy        127 LSSFLGTALAGIELDSLIKK----GDAAPDALAKASIDLINEIIGNLSQSTQTIEAFSSQ 182
              | :|  | ||:|:  || |::|         | : : :||:|:||:::     |    |:::|| |
Db        135 LQNFTGIALSGMALDELLRKQREGEDISQNDIAKSSIELINQLVDTVSSINSTVDSFSEQ 194

Qy        183 LAKLGSTISQAKGFSNIGNKLQNL-NFSKTNLGLEIITGLLSGISAGFALADKNASTGKK 241
              | :||| :|        |::| |||||  :         ||::::|:|| :|| |  | :| ||  |
Db        195 LNQLGSFLSSKPRLSSVGGKLQNLPDLGPLGDGLDVVSGILSAVSASFILGNSDAHTGTK 254

Qy        242 VAAGFELSNQVIGNVTKAISSYVLAQRVAAGLSTTGAVAALITSSIMLAISPLAFMNAAD 301
               ||| ||: ||:||| ||:| |:||||:| |||||  | |  ||||::||||||||:|: |||
Db        255 AAAGIELTTQVLGNVGKAVSQYILAQRMAQGLSTTAASAGLITSAVMLAISPLSFLAAAD 314

Qy        302 KFNHANALDEFAKQFRKFGYDGDHLLAEYQRGVGTIEASLTTISTALGAVSAGVSAAAVG 361
               ||  |  |: ::::|:|  |:|| ||| : :  | |:| ||||:| | :||||||||:
Db        315 KFERAKQLESYSERFKKLNYEGDALLAGFHKETGAIDAGLTTINTVLSSVSAGVSAASSA 374

Qy        362 SAVGTPIALLVAGVTGLISGILEASKQAMFESVANRLQGKILEWEKQNGGQNYFDKGYDS 421
               | :| ||::||: :|| ||||||||||||| ||  :    :|  ||||:: |:||: |||:
Db        375 SLIGAPISMLVSALTGTISGILEASKQAMFEHVAEKFAARINEWEKEH-GKNYFENGYDA 433

Qy        422 RYAAYLANNLKFLSELNKELEAERVIAITQQRWDNNIGELAGITKLGERIKSGKAYADAF 481
              |:||:| ::|  |:: :::   || :||| | ||   |||||||:  :| :||| | :
Db        434 RHAAFLEDSLSLLADFSRQHAVERAVAITHQHWDEKIGELAGITRNADRSQSGKPYINYL 493

Qy        482 EDGKKVEAG----SNITLDAKTGIIDISNSNGKKTQALHFTSPLLTAGTESRERLTNGKY 537
              |:|  :||      :       | : | ||:  | |    :  | | :|  | | | ||| :|||
Db        494 ENGGLLEAQPKEFTQQVFDPQKGTIDL--STGNVSSVLTFITPTFTPGEEVRERKQSGKY 551

Qy        538 SYINKLKFGRVKNWQVTDGEASSK--LDFSKVIQRVAETEGTDEIGLIVNAKAGNDDIFV 595
               |:  |          || |   |  : |    |:||:|| |  :      :  :|      :| ::
Db        552 EYMTSLIVNGKDTWSV-KGIKNHKGVYDYSKLIQFVEKNTKHYQARIISELGDKDDVVYS 610

Qy        596 GQGKMNIDGGDGHDRVFYSKDGGFGNITVDGTSATEAGSYTVNRKVARGD--IYHEVVKR 653
               | |   :  |:|:| | | |:|      | :|:| | |  : | | | | : :  ||  :  ||||
Db        611 GAGSSEVFAGEGYDPVSYNKT-DVGKLTIDATGAPKPGEYIVPKNM-YGDVEVLQEVVKE 668

Qy        654 QETKVGKRTETIQYRDYELRKVGYGYQSTDNLKSVEEVIGSQFNDVFKGSKFNDIFHSG- 712
              || |||||| |||||:| |  |   |      ||| ||||:|| : :| ||| ||||||
Db        669 QEVSVGKRTEKIQYRDFEFRTGGIPYDVIDNLHSVEELIGGKHDDEFKGGKFNDIFHGAD 728

Qy        713 -----------------EGDDLLDGGAGDDRLFGGKGNDRLSGDEGD------------- 742
                                :||| : || |||:|||| |||:||| :|:
Db        729 GNDYIEGNYGNDRLYGDDGDDYISGGQGDDQLFGGSGNDKLSGGDGNNYLTGGSGNDELQ 788

Qy        743 ----------------------DLLDGGSGDDVLNGGAGNDVYIFRKGDGNDTLYD-GT 778
                                    ||||| |:| |||| |||:|:: :  |: |: | |
Db        789 AHGAYNILSGGTGDDKLYGGGGIDLLDGGEGNDYLNGGFGNDIYVYGQNYGHHTIADEGG 848

Qy        779 GNDKLAFADANISDIMIERTKEGIIVKR-----------NDHSGSINIPRWYITSNLQNY 827
                 |:|  :| :  ||  :|     :|: :            || :| |    |:
Db        849 KGDRLHLSDISFDDIAFKRVGNDLIMNKAINGVLSFNESNDVNG-ITFKNWFAKD----- 902

Qy        828 QSNKTDHKIEQLIGKDGSYITSDQIDKILQDKKDGTVITSQELKKLADENKSQKLSASDI 887
                |    :| :| :  |||  |    ::|||  :   :      | :|     ::   :: : :
Db        903 ASGADNHLVEVITDKDGREI---KVDKIPHNNNE----RSGYIKASNIASEKNMVNITSV 955

Qy        888 ASSLNKLVGSMALFGTAN 905
              |: :||:: |:: | : :
Db        956 ANDINKIISSVSGFDSGD 973
```

```
RESULT 35
Q9L469
ID   Q9L469      PRELIMINARY;      PRT;  1706 AA.
AC   Q9L469;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Bifunctional hemolysin-adenylate cyclase (EC 4.6.1.1).
GN   CYAA.
OS   Bordetella parapertussis.
OC   Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC   Alcaligenaceae; Bordetella.
OX   NCBI_TaxID=519;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=63.2;
RA   Boursaux-Eude C., Guiso N.;
RT   "Sequencing of the Bordetella parapertussis adenylate cyclase-
RT   hemolysin-encoding gene.";
RL   Submitted (SEP-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AJ249835; CAB76450.1; -.
DR   InterPro; IPR005165; Anthrax_toxA.
DR   InterPro; IPR001343; Hemlysn_Ca_bind.
DR   InterPro; IPR003995; RtxA.
DR   Pfam; PF03497; Anthrax_toxA; 1.
DR   Pfam; PF00353; hemolysinCabind; 17.
```

```
DR   Pfam; PF02382; RTX; 1.
DR   PRINTS; PR00313; CABNDNGRPT.
DR   PRINTS; PR01488; RTXTOXINA.
DR   PROSITE; PS00330; HEMOLYSIN_CALCIUM; 3.
KW   Lyase.
SQ   SEQUENCE   1706 AA;  177038 MW;  71750E2D0BBF3B64 CRC64;

  Query Match             17.2%;  Score 799;  DB 2;  Length 1706;
  Best Local Similarity   27.2%;  Pred. No. 2.3e-27;
  Matches  247;  Conservative  143;  Mismatches  347;  Indels  170;  Gaps   27;

Qy        106 GLDSVENIDRKLGKASNVLSTLSSFLGTALAGIELDSLI-----KKGDAAP--------- 151

Db        403 GYDSLDGV----GSRSFSLGEVSDMAAVEAAELEMTRQVLHAGARQDDAEPGVSGASAHW 458

Qy        152 --------DALAKA-----SIDLINEI----IGNLSQSTQTIEAFSSQLAKLGSTISQ-A 193

Db        459 GQRALQGAQAVAAAQRLVHAIALMTQFGRAGSTNTPQEAASLSAAVFGLGEASSAVAETV 518

Qy        194 KGFSNIGNKLQNLNFSKTNLGLEIITGLLSGISAGFALADKNASTGKKVAAGFELSNQVI 253

Db        519 SGFFR-GSSRWAGGFGVAGGAMALGGGIAAAVGAGMSLTD-DAPAGQKAAVGAEIALQLT 576

Qy        254 GNVTKAISSYVL----AQRVAAGLSTTGAVAALITSSIMLAISPLAFMNAADKFNHANAL 309

Db        577 GGTVELASSIALALAAARGVTSGLQVAGASAGAAAGALAAALSPMEIYGLVQQSHYADQL 636

Qy        310 DEFAKQFRKFGYDGDHLLAEYQRGVGTIEASLTTISTALGAVSAGVSAAAVGSAVGTPIA 369

Db        637 DKLAQESSAYGYEGDALLAQLYRDKTAAEGAVAGVSAVLSTVGAAVSIAAAASVVGAPVA 696

Qy        370 LLVAGVTGLISGILEASKQAMFESVANRLQGKILEWEKQNGGQNYFDKGYDSRYAAYLAN 429

Db        697 VVTSLLTGALNGILRGVQQPIIEKLANDYARKI---DELGGPQAYFEKNLQARH-EQLAN 752

Qy        430 N---LKFLSELNKELEAERVIAITQQRWDNNIGELAGITKLGERIKSGKAYADAFEDGKK 486

Db        753 SDGLRKMLADLQAGWNASSVIGVQTTEISKSALELAAITGNADNLKSADVFVDRFIQGER 812

Qy        487 VEAGSNITLDAKTGIIDISNSNGKKTQALHFTSPLLTAGTESRERLTNGKYSYINKLKF- 545

Db        813 V-AGQPVVLDVAAGGIDIASRKGER-PALTFITPLAAPGEEQRRRTKTGKSEFTTFVEIV 870

Qy        546 GRVKNWQVTDGEASSKLDFSKVIQRVAETEGT--DEIGLIVNAKAGNDDIFVGQGKMNID 603

Db        871 GKQDRWRIRDGAADTTIDLAKVVSQLVDANGVLKHSIKLEVIGGDGDDVVLANASRIHYD 930

Qy        604 GGDGHDRVFYSKDGGFGNITVDGTSATEAGSYTVNRKVARGDIYHEVVKRQETKVGKRTE 663

Db        931 GGAGTNTVSYAALGRQDSITV----SADGERFNVRKQLNNANVYREGVATQKTAYGKRTE 986

Qy        664 TIQYRDYELRKVGYGYQSTDNLKSVEEVIGSQFNDVFKGSKFNDIFHSGEGDDLLDGGAG 723

Db        987 NVQYRHVELARVGQ-LVEVDTLEHVQHIIGGAGNDSITGNAHDNFLAGGAGDDRLDGGAG 1045

Qy        724 DDRLFGGKGNDRLSGDEGDDL----------LDGGSG--------------------- 750

Db       1046 NDTLVGGEGHNTVVGGAGDDVFLQDLGVWSNQLDGGAGVDTVKYNVHQPSEERLERMGDT 1105

Qy        751 --------------------------------------------DDVLNGGAGNDVY-- 763

Db       1106 GIHADLQKGTVEKWPALNLFSVDHVKNIENLHGSSLNDSIAGDDRDNELWGDDGNDTIHG 1165

Qy        764 -----IFRKGDGNDTLYDGTGNDKLAFADANISDIM-----IERTKEGIIVKRNDHSGSI 813

Db       1166 RGGDDILRGGLGLDTLYGEDGNDIFLQDDETVSDDIDGGAGLDTVDYSAMI----HAGKI 1221

Qy        814 NIPRWY---ITSNLQNYQSNKTDHK---------IEQLIG---KDGSYITSDQIDKILQ 857

Db       1222 VAPHEYGFGIEADLSEGWVRKAARRGMGYYDSVRSVENVIGTSMKD-VLIGDAQANTLMG 1280

Qy        858 DKKDGTV 864

Db       1281 QGGDDTV 1287


RESULT 36
Q51865
ID   Q51865      PRELIMINARY;      PRT;   233 AA.
AC   Q51865; Q51867;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Leukotoxin A (Fragment).
OS   Pasteurella haemolytica.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC   Pasteurellaceae; Mannheimia.
OX   NCBI_TaxID=75985;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=T10 SEROTYPE AND T3 SEROTYPE;
RA   Lainson A.F., Aitchison K.D., Donachie W.;
RL   Submitted (JUN-1993) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; Z22887; CAA80501.1; -.
DR   EMBL; Z22884; CAA80498.1; -.
DR   HSSP; Q03023; 1KAP.
DR   InterPro; IPR001343; Hemlysn_Ca_bind.
DR   Pfam; PF00353; hemolysinCabind; 4.
DR   PRINTS; PR00313; CABNDNGRPT.
FT   NON_TER       1      1
SQ   SEQUENCE   233 AA;  25345 MW;  AD6C8408731C3F99 CRC64;

  Query Match             10.4%;  Score 482;  DB 2;  Length 233;
  Best Local Similarity   41.5%;  Pred. No. 2.1e-14;
  Matches   97;  Conservative   53;  Mismatches   70;  Indels   14;  Gaps    8;

Qy        698 DVFKGSKFNDIFHSGEGDDLLDGGAGDDRLFGGKGNDRLSGDEGDDLLDGGSGDDVLNGG 757

Db          1 DIFKGSKFNDAFNGGDGVDTIDGNDGNDRLFGGKGDDIIDGGNGDDFIDGGKGNDLLHGG 60

Qy        758 AGNDVYIFRKGDGNDTLYDGTGNDKLAFADANISDIMIERTKEGIIVKRNDHSGSINIPR 817

Db         61 KGDDIFVHRQGDGNDSITESEGNDKLSFSDSNLKDLTFEKVNHHLVI-TNTKQEKVTIQN 119

Qy        818 WY----ITSNLQNYQSNKTDHKIEQLIGKDGSYITSDQIDKILQDKKDGTVITSQELKKL 873

Db        120 WFREAEFAKTIQNYVATR-DDKIEEIIGQNGERITSKQVDELIE--KGNGKIAQSELTKV 176

Qy        874 ADENKSQKLSASDIASSLNKLVGSMALFGTANSVSSNALQPITQPTQGILAPSV 927

Db        177 VDNYQLLKYS-RDASNSLDKLISSASAFTSSND-SRNVL---ASPT-SMLDPSL 224


RESULT 37
Q9PEL7
ID   Q9PEL7      PRELIMINARY;      PRT;  1636 AA.
AC   Q9PEL7;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Hemolysin-type calcium binding protein.
GN   XF1011.
OS   Xylella fastidiosa.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC   Xanthomonadaceae; Xylella.
OX   NCBI_TaxID=2371;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=9a5c;
RX   MEDLINE=20365717; PubMed=10910347;
RA   Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,
RA   Alvarenga R., Alves L.M.C., Araya J.E., Baia G.S., Baptista C.S.,
RA   Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Briones M.R.S.,
RA   Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carrer H.,
RA   Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,
RA   Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorry H.,
```

```
RA   Facincani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,
RA   Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furlan L.R.,
RA   Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,
RA   Ho P.L., Hoheisel J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,
RA   Krieger J.E., Kuramae E.E., Laigret F., Lambais M.R., Leite L.C.C.,
RA   Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,
RA   Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Marino C.L.,
RA   Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,
RA   Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,
RA   Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,
RA   Nhani A. Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,
RA   de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,
RA   Peixoto B.R., Pereira G.A.G., Pereira H.A. Jr., Pesquero J.B.,
RA   Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,
RA   de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,
RA   da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,
RA   da Silveira J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,
RA   de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tsuhako M.H.,
RA   Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,
RA   Zago M.A., Zatz M., Meidanis J., Setubal J.C.;
RT   "The genome sequence of the plant pathogen Xylella fastidiosa.";
RL   Nature 406:151-159(2000).
DR   EMBL; AE003938; AAF83821.1; -.
DR   InterPro; IPR001343; Hemlysn_Ca_bind.
DR   InterPro; IPR003995; RtxA.
DR   Pfam; PF00353; hemolysinCabind; 18.
DR   PRINTS; PR00313; CABNDNGRPT.
DR   PRINTS; PR01488; RTXTOXINA.
DR   PROSITE; PS00330; HEMOLYSIN_CALCIUM; 4.
KW   Complete proteome.
SQ   SEQUENCE   1636 AA;  173012 MW;  76A6E54271859CB2 CRC64;

  Query Match             8.9%;  Score 415;  DB 16;  Length 1636;
  Best Local Similarity   22.4%;  Pred. No. 3.2e-10;
  Matches  226;  Conservative  150;  Mismatches  323;  Indels  308;  Gaps   48;

Qy         25 LYLAIPKDYDPQKGGTLNDFIKAADELG---IARLAEEPNHTETAKKSVDTVNQFLSLTQ 81
              :|||:|      |    | : | :|  :|      ||  ||  |:           | ||
Db        793 MYLALP---PRQLSATFDAFAEAYAKLKESIYIRLVLEPRLTDYLSG--------LRLTD 841

Qy         82 TG--IAISATKLEKFLQKHSTNKLAKGLDSVENIDRKLGKASNVLSTLSSFLGTALAGIE 139
              :|  |:  ||  |:   :  |: |  | :|    || :             | :| :
Db        842 NNGVMAWDASGLEAKLDQTWQHNKAQALQDVMDLYR---YGSNAV---------AASGWK 889

Qy        140 ----LDSLIKKGDAAPD---ALAKASIDLINEIIGNLSQSTQTIEAFSSQLAKLGSTISQ 192
                  | :| :  | ||    ||| | | |::  ||    |    :|        |: : :
Db        890 PFDALRHMIDRTAATPDGRQALAAAEITLVS---GNAEGS----DAADLLFGDAGANLLR 942

Qy        193 AKGFSNIGNKLQNLNFSKTNLGLEIITG-----LLSGISAGFALADKNASTGKKVAAGFE 247
                 |                   | ::::|    | | :    |      |  |  | |
Db        943 GGG------------------GDDVLSGGGGNDTLEGGAGNDTL---YGDDGDDVLDGGE 981

Qy        248 LSNQVIGNVTKAI--SSYVLAQRVAAGLSTTGAVAALITSSIMLAISPLAFMNAADKFNH 305
              ||::  |    :   ||  |  |   |    |     :   |   | :     |  |  |
Db        982 GSNRLEGGAGNDVLKVSYWSADNVLIG----GTGDDTLYGS---AFADTYLFNQGD--GH 1032

Qy        306 ANALDEFAKQFRKFG----------YDGDHLLAEYQRGVGTI--------------EAS 340
                 :::             ||     :|  :: :   |  :|                 |
Db       1033 DTIIEQGGTDKLVFGEGILAADVRLLREGQDVVLDLGNGHDSIRLKDWLTSNGTRNHSAD 1092

Qy        341 LTTISTALGAVSAGVSAAAVG-SAVGTPIALLVAGVTG---LISG----ILEASKQAMFE 392
                : |  |  | :    : :::| : :|||   :  |    |: |    :|: :
Db       1093 IEQIVFADGTLWTPETLSSMGLTTLGTPGNDTLKGWQGKDILLGGAGDDVLDGGEG---- 1148

Qy        393 SVANRLQG-------KILEWEKQN---GGQNYFDKGYDSRYAAYLANNLKFLSELNKELE 442
               :|||:|        |:  |   |   ||      | |: | :   |:   |
Db       1149 --SNRLEGGAGDDVLKVSYWSADNVLSGG-----TGDDTLYGSAFADTYLF--------- 1192

Qy        443 AERVIAITQQRWDNNIGELAGITKL--GERIKSGKAYADAFEDGKKVEAGSNITLDAKTG 500
                         |     :  |  |  |   ||  || |  :               | : :   |  : : ||
Db       1193 -------NQGDGHDTIIEQGGTDKLVFGEGILAA--------DVRLLREGQDVVLD---- 1233

Qy        501 IIDISNSNGKKTQALHFTSPLLTAGTESRERLTNGKYSYINKLKFGRVKNWQVTDGEASS 560
                                            | ||  |            |:|:|  ::|  :
Db       1234 -----------------------------LGNGHDSI-------RLKDWLTSNGTRNH 1255

Qy        561 KLDFSKVIQRVAETEGT-------DEIGLIVNAKAGND--------DIFV-GQGKMNIDG 604
                 |     |:::   :||          :||        :|||        || : | |   :||
Db       1256 SAD----IEQIVFADGTLWTPETLSSMGLTTLGTSGNDTLKGWQGKDILLGGAGDDVLDG 1311

Qy        605 GDGHDRV------------FYSKD----GGFGNITVDGTSATEAGSYTVNRKVARGDIYH 648
              |:| :|:                    ::| |    || |: |: |::     |  :|  |   :|| :
Db       1312 GEGSNRLEGGAGDDVLKVSYWSADNVLIGGTGDDTLYGSAF--ADTYLFN----KGDGHD 1365

Qy        649 EVVKRQET------------------------KVGKRTETIQYRDYELRKVGYGYQSTD 683
              ::::  |                              :|   ::|: :|:  |     |      | |
Db       1366 TIIEQGGTDKLVFGEGILAADVRLLREGQDVVLDLGNGHDSIRLKDW-LTSNGTRNHSAD 1424

Qy        684 NLKSVEEVI-------------------GSQFNDVFKGSKFNDIFHSGEGDDLLDGGAG 723
                   :|:::                                   |:  ||   ||  :   ||      |:|||:||||  |
Db       1425 ----IEQIVFADGTLWTPETLSSMGLTTLGTPGNDTLKGWQGKDILLGGDGDDVLDGGEG 1480

Qy        724 DDRLFGGKGND--RLSGDEGDDLLDGGSGDDVLNGGAGNDVYIFRKGDGNDTLYDGTGND 781
               :|| ||  |:|    ::|       |::|   ||:|||   |   |   |   |:|  ||||:||: :   |  |
Db       1481 SNRLEGGAGDDVLKVSYWSADNVLSGGTGDDTLYGSAFADTYLFNKGDGHDTIIEQGGTD 1540

Qy        782 KLAF-ADANISDIMIERTKEGIIVKRNDHSGSINIPRWYITSNLQNYQSNKTDHKIEQLI 840
              || |  |   :    :       |: : : :      |          : :     |:          | : |::|  |   :
Db       1541 KLVFGAGLSRKEARFTRSGDDLSLLFNGSEDQVTVAGWF----------NGSGHQVESLV 1590

Qy        841 GKDGSYITSDQIDKILQDKKDGTVITSQELKKLADENKSQKLSASDI 887
                 :||: :  |  ::::::       :|:  :    :  :  :| :| |||
Db       1591 FQDGT-VLSAEVERLIAAMALSPAVTTMQ-ASVGETKESPRLVASSI 1635

RESULT 38
Q9PFI9
ID   Q9PFI9       PRELIMINARY;      PRT;  1208 AA.
AC   Q9PFI9;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Hemolysin-type calcium binding protein.
GN   XF0668.
OS   Xylella fastidiosa.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC   Xanthomonadaceae; Xylella.
OX   NCBI_TaxID=2371;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=9a5c;
RX   MEDLINE=20365717; PubMed=10910347;
RA   Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,
RA   Alvarenga R., Alves L.M.C., Araya J.E., Baia G.S., Baptista C.S.,
RA   Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Briones M.R.S.,
RA   Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carrer H.,
RA   Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,
RA   Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorry H.,
RA   Facincani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,
RA   Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furlan L.R.,
RA   Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,
RA   Ho P.L., Hoheisel J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,
RA   Krieger J.E., Kuramae E.E., Laigret F., Lambais M.R., Leite L.C.C.,
RA   Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,
RA   Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Marino C.L.,
RA   Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,
RA   Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,
RA   Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,
RA   Nhani A. Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,
RA   de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,
RA   Peixoto B.R., Pereira G.A.G., Pereira H.A. Jr., Pesquero J.B.,
RA   Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,
RA   de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,
RA   da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,
RA   da Silveira J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,
RA   de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tsuhako M.H.,
```

```
RA   Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,
RA   Zago M.A., Zatz M., Meidanis J., Setubal J.C.;
RT   "The genome sequence of the plant pathogen Xylella fastidiosa.";
RL   Nature 406:151-159(2000).
DR   EMBL; AE003910; AAF83478.1; -.
DR   InterPro; IPR001343; Hemlysn_Ca_bind.
DR   Pfam; PF00353; hemolysinCabind; 15.
DR   PRINTS; PR00313; CABNDNGRPT.
DR   PROSITE; PS00330; HEMOLYSIN_CALCIUM; 3.
KW   Complete proteome.
SQ   SEQUENCE   1208 AA;  128403 MW;  DF151ACD7741872B CRC64;

  Query Match              8.8%;  Score 411;  DB 16;  Length 1208;
  Best Local Similarity   22.0%;  Pred. No. 3.1e-10;
  Matches  227;  Conservative  169;  Mismatches  344;  Indels  294;  Gaps   46;

Qy         36 QKGGTLNDFIKAADELGIARLAEEPNHTETAKKSVDTVNQFL-----SLTQTG-----IA 85
              | | :  : :       |||| :    | :||      ||  | |       : |:|        :
Db        286 QDGQSQANELFTLSSLGIASITLTPTNTE----DVDLGNGNLIDNRGTYTRTDGRTGVVG 341

Qy         86 ISATKLEKFLQKHS-------TNKLAKGLDSVENID--RKLGKASNVLSTLSSFLGTALA 136
                     || | : :|             :    |: | ::          | | :|:::    | : :
Db        342 DLQLGLEHFYRDYSGAHDKVTVSDAARALPAISGSGAVRDLREAASLSPALLAAVQALTP 401

Qy        137 GIELDSLIKKGDAAPDALAKASIDLINEIIGNLS------------------------ 170
              |    |::                :|::|  :    :     |
Db        402 GTTRDTM-----------RAALDPLVALWAGTSAMPSTEQRLETSAAVPRTVYYHGAVP 449

Qy        171 -----QSTQTIEAF-SSQLAKLGSTISQAKGFSNIGNKLQNLNFSKTNLGLEIITGLLSG 224
                   |  : ::|:    | |:||   |:  :  |:    |:  |  :       :  | |   |
Db        450 AAVTAQGKEAVQAWIQQQHARLGPIIAILEKFN--GSSLVSERNGQVSTGGETFTW---- 503

Qy        225 ISAGFALADKNASTGKKVAAGFELSNQVI---GNVTKAISSYVLAQRVAAGLSTTGAVAA 281
                                         |:||    |:     :|   :  ::: : :|     ::
Db        504 ------------------------NRVIHPDGHSEDVMSILLQPEQINSLMSAYASLKE 538

Qy        282 LITSSIMLA------ISPLAFMNAADKFN-HANALDEFAKQFRKFGYDGDHLLAE----Y 330
                  : ::|:          :|  |       : |  |    |:||:   ||       : ::      |   :     |
Db        539 FAYAGLVLSTRLSDYLSGLTLTYSNDAFGWDASALE--AKLDHTWQHNKTQALQDVMDLY 596

Qy        331 QRGVGTIEASLTTISTALGAVSAGVSAAAVGSAVGTPIALLVAGVTGLISGILEASKQA- 389
              : |  :  ||              ||    :        :        ||  |  : | |:||   | |   |
Db        597 RYGSNAVAASGWKPFDALRHMIDRTAATPDGRQ-----ALAAAEIT-LVSGSAEGSTAAD 650

Qy        390 --MFESVANRLQGKILEWEKQNGGQN------------YFDKGYDSRYAAYLANNLKFL 434
                :: || |:|    :  :   :||                        | | | |:  |        |    |
Db        651 LIFGDAGANILRGGAGD-DVLSGGDGDDTLYGGDGNDTLYGDAGNDTLYGG--GGNDLLL 707

Qy        435 ---------SELNKELE---AERVIAITQQRWDNNIGELAGITKLGERIKSGKAYADAF 481
                         | :  ||     :  |: ::       || :           |        |:      | |:|| :
Db        708 GGDGDDVLDGGEGSNRLEGGAGDDVLKVSYWSADNVL-----IGGTGDDTLYGSAFADTY 762

Qy        482 ----EDGKK--VEAGSNITLDAKTGII--DISNSNGKKTQALHFTSPLLTAGTESRERLT 533
                       ||        :|  |         |     ||:    |:                          || |   :     |
Db        763 LFNQGDGHDTIIEQGGTDKLVFGEGIVAADVR---------------LLRQGQDVVLDLG 807

Qy        534 NGKYSYINKLKFGRVKNWQVTDGEASSKLDFSKVIQRVAETEGT-------DEIGLIVNA 586
              || |          |:|:|    ::|    :     |       |::: :  :||                  :||
Db        808 NGHDSI-------RLKDWLTSNGTRNHSAD----IEQIVFADGTLWTPETLSSMGLTTLG 856

Qy        587 KAGND--------DIFV-GQGKMNIDGGDGHDRV-----------FYSKD----GGFGN 621
                :|||              || :  | |      :|||:|  :|:                  ::|  |      || |:
Db        857 TSGNDTLKGWQGKDILLGGAGDDVLDGGEGSNRLEGGAGNDVLKVSYWSADNVLIGGTGD 916

Qy        622 ITVDGTSATEAGSYTVNRKVARGDIYHEVVKRQET------------------------- 656
                |: |::       | :| |      :||  :   ::::   |
Db        917 DTLYGSAF--ADTYLFN----QGDGHDTIIEQGGTDKLVFGEGILAADVRLLREGQDVVL 970

Qy        657 KVGKRTETIQYRDYELRKVGYGYQSTDNLKSVEEVI--------------------GSQF 696
                :|     ::|: :|: :|    |      |    | |        :|:::                          |:
Db        971 DLGNGHDSIRLKDW-LTSNGTRNHSAD----IEQIVFADGTLWTPETLSSMGLTTLGTPG 1025

Qy        697 NDVFKGSKFNDIFHSGEGDDLLDGGAGDDRLFGGKGND--RLSGDEGDDLLDGGSGDDVL 754
                  || || :   ||    |:|||:||||  |  :|  ||  |||    ::|       |::|  ||:||| |
Db       1026 NDTLKGWQGKDILLGGDGDDVLDGGEGSNRLEGGAGNDVLKVSYWSADNVLIGGTGDDTL 1085

Qy        755 NGGAGNDVYIFRKGDGNDTLYDGTGNDKLAFADA-NISDIMIERTKEGIIVKRNDHSGSI 813
                  |  |    | |:|  ||||:||: : :| |||  | :    :    :        |: : : :    |          :
Db       1086 YGSAFADTYLFNKGDGHDTIIEQSGTDKLVFGEGLHQKEARFTRSGDDLSLLFNGSEDQV 1145

Qy        814 NIPRWYITSNLQNYQSNKTDHKIEQLIGKDGSYITSDQIDKILQDKKDGTVITSQELKKL 873
                  |  |:                        |  :  |::|  |:  :||:  :  |  ::::::            |:  :    :
Db       1146 TIAGWF----------NGSGHQVESLVFQDGT-VLSGEVERLIAAMALSPAATTLQ-ASV 1193

Qy        874 ADENKSQKLSASDI 887
                  |  :|  :|  || |
Db       1194 GDHKESPRLVASSI 1207


RESULT 39
Q9PAT8
ID   Q9PAT8      PRELIMINARY;      PRT;  2064 AA.
AC   Q9PAT8;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Bacteriocin.
GN   XF2407.
OS   Xylella fastidiosa.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC   Xanthomonadaceae; Xylella.
OX   NCBI_TaxID=2371;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=9a5c;
RX   MEDLINE=20365717; PubMed=10910347;
RA   Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,
RA   Alvarenga R., Alves L.M.C., Araya J.E., Baia G.S., Baptista C.S.,
RA   Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Briones M.R.S.,
RA   Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carrer H.,
RA   Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,
RA   Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorry H.,
RA   Facincani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,
RA   Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furlan L.R.,
RA   Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,
RA   Ho P.L., Hoheisel J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,
RA   Krieger J.E., Kuramae E.E., Laigret F., Lambais M.R., Leite L.C.C.,
RA   Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,
RA   Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Marino C.L.,
RA   Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,
RA   Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,
RA   Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,
RA   Nhani A. Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,
RA   de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,
RA   Peixoto B.R., Pereira G.A.G., Pereira H.A. Jr., Pesquero J.B.,
RA   Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,
RA   de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,
RA   da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,
RA   da Silveira J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,
RA   de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tsuhako M.H.,
RA   Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,
RA   Zago M.A., Zatz M., Meidanis J., Setubal J.C.;
RT   "The genome sequence of the plant pathogen Xylella fastidiosa.";
RL   Nature 406:151-159(2000).
DR   EMBL; AE004049; AAF85206.1; -.
DR   InterPro; IPR001343; Hemlysn_Ca_bind.
DR   Pfam; PF00353; hemolysinCabind; 31.
DR   PRINTS; PR00313; CABNDNGRPT.
DR   PROSITE; PS00330; HEMOLYSIN_CALCIUM; 7.
KW   Complete proteome.
SQ   SEQUENCE   2064 AA;  218776 MW;  6265655E90DA1D9A CRC64;

  Query Match              8.4%;  Score 391;  DB 16;  Length 2064;
  Best Local Similarity   25.4%;  Pred. No. 5.2e-09;
  Matches  135;  Conservative   89;  Mismatches  163;  Indels  144;  Gaps   23;
```

```
Qy        446 VIAITQQRWDNNIGELAGITKLGERIKSGKAYADAF----EDGKK--VEAGSNITLDAKT 499
              |: :  :  || :     |    |:     | ||||  :     ||     :| | : ||
Db       1588 VLKVAYRSADNVL-----IGGTGDDTLYGSAYADTYLFNKGDGHDTIIEQGGDDTLVFGA 1642

Qy        500 GIIDISNSNGKKTQALHFTSPLLTAGTESRERLTNGKYSYINKLKFGRVKNWQVTDGEAS 559
              ||:          :|        :|  |  :    | ||  |         |:|:| :|| :
Db       1643 GIV--------ASQV-----RVLREGQDVVLDLGNGHDSI-------RLKDWLTSDGYRN 1682

Qy        560 SKLDFSKVIQRVAETEGT-------DEIGLIVNAKAGNDDIFVGQGKMNIDGGDGHDRVF 612
                 |    |:::  :||          :||      :|||  :   |||  : |||| | :
Db       1683 GNND----IEQIVFADGTIWTPETLSSMGLTTLGTSGNDTLKGWQGKDILLGGDGDDVL- 1737

Qy        613 YSKDGGFG---------------------NITVDGT------SATEAGSYTVNRKVARGD 645
                 ||| |                    |: : ||       :  | :|  |     :||
Db       1738 ---DGGMGSNRLEGGAGNDVLKVAYRSADNVLIGGTGDDTLYGSGYADTYLFN----KGD 1790

Qy        646 IYHEVVKR--QETKV----------------------GKRTETIQYRDYELRKVGYGYQ 680
               :  ::::   :| |                      |    ::|: :|:      ||:
Db       1791 GHDTIIEQGGDDTLVFGAGIVASQVRVLREGQDVVLDLGNGHDSIRLKDW---LTSDGYR 1847

Qy        681 STDNLKSVEEVI--------------------GSQFNDVFKGSKFNDIFHSGEGDDLLDG 720
               : :|    :|:::                    |:  ||  || :  ||    |:|||:|||
Db       1848 NGNN--DIEQIVFADGTIWTPETLSSMGLTTLGTSGNDTLKGWQGKDILLGGDGDDVLDG 1905

Qy        721 GAGDDRLFGGKGND--RLSGDEGDDLLDGGSGDDVLNGGAGNDVYIFRKGDGNDTLYDGT 778
              | | :|| || |||  :::    |::| ||:|||  |  |     | |:| ||||:||: :
Db       1906 GMGSNRLEGGAGNDVLKVAYRSADNVLIGGTGDDTLYGSAYADTYLFNKGDGHDTIIEQG 1965

Qy        779 GNDKLAFADANISDIMIERTKEG--IIVKRNDHSGSINIPRWYITSNLQNYQSNKTDHKI 836
              |:| | |  | :          || |   : :   |      :  |  |:            | : |::
Db       1966 GDDTLVFG-AGLHQKEARFTKSGNDLSILFNASEDQVTIAGWF---------NGSGHQV 2014

Qy        837 EQLIGKDGSYITSDQIDKILQDKKDGTVITSQELKKLADENKSQKLSASDI 887
              | |: :||: : | :::::       : :|: :   :  |   :| :| || |
Db       2015 ESLVFQDGT-VLSGEVERLIAAMALSSAVTTTQ-ASVRDTKESHRLVASSI 2063
```

```
RESULT 40
Q51868
ID   Q51868      PRELIMINARY;      PRT;   208 AA.
AC   Q51868; Q51866;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Leukotoxin A (Fragment).
OS   Pasteurella haemolytica.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC   Pasteurellaceae; Mannheimia.
OX   NCBI_TaxID=75985;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=T4 SEROTYPE AND T15 SEROTYPE;
RA   Lainson A.F., Aitchison K.D., Donachie W.;
RL   Submitted (JUN-1993) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; Z22885; CAA80499.1; -.
DR   EMBL; Z22886; CAA80500.1; -.
DR   HSSP; Q03023; 1AKL.
DR   InterPro; IPR001343; Hemlysn_Ca_bind.
DR   Pfam; PF00353; hemolysinCabind; 3.
FT   NON_TER       1      1
SQ   SEQUENCE   208 AA;  22803 MW;  25301410C85A4CC5 CRC64;

  Query Match             8.2%;  Score 383;  DB 2;  Length 208;
  Best Local Similarity   37.8%;  Pred. No. 4.7e-10;
  Matches   79;  Conservative   52;  Mismatches   64;  Indels   14;  Gaps    8;

Qy        723 GDDRLFGGKGNDRLSGDEGDDLLDGGSGDDVLNGGAGNDVYIFRKGDGNDTLYDGTGNDK 782
              |:|||||||||:| : |  ||| :||| |:|:|:|| |:|::: |:|||||:: : ||||
Db          1 GNDRLFGGKGDDIIDGGNGDDFIDGGKGNDLLHGGKGDDIFVHRQGDGNDSITESEGNDK 60

Qy        783 LAFADANISDIMIERTKEGIIVKRNDHSGSINIPRWY----ITSNLQNYQSNKTDHKIEQ 838
              |:|:|:|: |: |:    :::  |      : | |:            ::|| : : | |||:
Db         61 LSFSDSNLKDLTFEKVNHHLVI-TNTKQEKVTIQNWFREAEFAKTIRNYVATR-DDKIEE 118

Qy        839 LIGKDGSYITSDQIDKILQDKKDGTVITSQELKKLADENKSQKLSASDIASSLNKLVGSM 898
              :||::|  ||| |:|::::  |    |    :| :: |  :   | |  | ::||:||: |
Db        119 IIGQNGERITSKQVDELIEKGKG--KIDKSDLSQVVDNYQLLKYS-RDASNSLDKLISSA 175

Qy        899 ALFGTANSVSSNALQPITQPTQGILAPSV 927
              : | ::|  | | |     ||  :| ||:
Db        176 SAFTSSND-SRNVL---ASPT-SMLDPSL 199

Search completed: February 17, 2004, 10:11:03
Job time : 57 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: February 17, 2004, 10:18:40 ; Search time 21 Seconds
(without alignments)
1867.722 Million cell updates/sec

Title: US-10-069-799-5
Perfect score: 927
Sequence: 1 MSNINVIKSNIQAGLNSTKS..........SSNALQPITQPTQGILAPSV 927

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 328717 seqs, 42310858 residues

Word size : 0

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	14	1.5	758	1	US-08-258-188-2	Sequence 2, Appli
2	14	1.5	758	1	US-08-526-813-2	Sequence 2, Appli
3	14	1.5	758	5	PCT-US95-08554-2	Sequence 2, Appli
4	14	1.5	1049	3	US-08-772-270A-11	Sequence 11, Appl
5	14	1.5	1053	4	US-09-062-126-10	Sequence 10, Appl
6	14	1.5	1244	5	PCT-US93-10500-2	Sequence 2, Appli
7	11	1.2	956	3	US-08-772-270A-8	Sequence 8, Appli
8	11	1.2	956	4	US-09-062-126-8	Sequence 8, Appli
9	11	1.2	1013	4	US-09-062-126-3	Sequence 3, Appli
10	11	1.2	1022	3	US-08-772-270A-2	Sequence 2, Appli
11	10	1.1	149	2	US-08-387-942C-45	Sequence 45, Appl
12	10	1.1	490	4	US-09-252-149B-26	Sequence 26, Appl
13	10	1.1	544	1	US-08-387-156-10	Sequence 10, Appl
14	10	1.1	544	2	US-08-694-865-10	Sequence 10, Appl
15	10	1.1	544	2	US-08-878-748-10	Sequence 10, Appl
16	10	1.1	544	3	US-09-124-491-10	Sequence 10, Appl
17	10	1.1	544	4	US-09-383-912-10	Sequence 10, Appl
18	10	1.1	699	2	US-08-694-865-16	Sequence 16, Appl
19	10	1.1	699	3	US-09-124-491-16	Sequence 16, Appl
20	10	1.1	699	4	US-09-383-912-16	Sequence 16, Appl
21	10	1.1	924	3	US-08-619-812-8	Sequence 8, Appli
22	10	1.1	926	1	US-07-908-253-2	Sequence 2, Appli
23	10	1.1	926	1	US-08-455-970A-2	Sequence 2, Appli
24	10	1.1	926	1	US-08-387-156-6	Sequence 6, Appli
25	10	1.1	926	2	US-08-694-865-6	Sequence 6, Appli
26	10	1.1	926	2	US-08-878-748-6	Sequence 6, Appli
27	10	1.1	926	2	US-08-535-837-2	Sequence 2, Appli
28	10	1.1	926	3	US-09-124-491-6	Sequence 6, Appli
29	10	1.1	926	4	US-09-383-912-6	Sequence 6, Appli
30	10	1.1	926	6	5476657-3	Patent No. 5476657
31	10	1.1	934	1	US-08-215-805A-80	Sequence 80, Appl
32	10	1.1	936	1	US-08-455-970A-12	Sequence 12, Appl
33	10	1.1	943	1	US-08-455-970A-10	Sequence 10, Appl
34	10	1.1	951	1	US-08-455-970A-14	Sequence 14, Appl
35	10	1.1	977	1	US-08-387-156-8	Sequence 8, Appli
36	10	1.1	977	2	US-08-694-865-8	Sequence 8, Appli
37	10	1.1	977	2	US-08-878-748-8	Sequence 8, Appli
38	10	1.1	977	3	US-09-124-491-8	Sequence 8, Appli
39	10	1.1	977	4	US-09-383-912-8	Sequence 8, Appli
40	10	1.1	1069	1	US-07-777-715-9	Sequence 9, Appli
41	10	1.1	1069	1	US-08-170-126-4	Sequence 4, Appli
42	10	1.1	1069	3	US-08-954-418-4	Sequence 4, Appli
43	10	1.1	1098	1	US-07-777-715-7	Sequence 7, Appli
44	10	1.1	1098	1	US-08-170-126-2	Sequence 2, Appli
45	10	1.1	1098	3	US-08-954-418-2	Sequence 2, Appli
46	10	1.1	1334	6	5476657-1	Patent No. 5476657
47	10	1.1	1403	1	US-07-908-253-3	Sequence 3, Appli
48	10	1.1	1403	2	US-08-694-865-17	Sequence 17, Appl
49	10	1.1	1403	2	US-08-387-942C-3	Sequence 3, Appli
50	10	1.1	1403	2	US-08-535-837-3	Sequence 3, Appli
51	10	1.1	1403	3	US-09-124-491-17	Sequence 17, Appl
52	10	1.1	1403	4	US-09-383-912-17	Sequence 17, Appl
53	8	0.9	128	4	US-09-328-352-5276	Sequence 5276, Ap
54	8	0.9	142	2	US-08-387-942C-41	Sequence 41, Appl
55	8	0.9	151	2	US-08-387-942C-46	Sequence 46, Appl
56	8	0.9	151	2	US-08-387-942C-50	Sequence 50, Appl
57	8	0.9	153	2	US-08-387-942C-47	Sequence 47, Appl
58	8	0.9	153	2	US-08-387-942C-51	Sequence 51, Appl
59	8	0.9	355	4	US-09-328-352-4560	Sequence 4560, Ap
60	8	0.9	499	4	US-09-328-352-5089	Sequence 5089, Ap
61	8	0.9	553	2	US-08-387-942C-2	Sequence 2, Appli
62	8	0.9	805	4	US-09-134-001C-4821	Sequence 4821, Ap
63	8	0.9	872	2	US-08-387-942C-5	Sequence 5, Appli
64	8	0.9	997	2	US-08-387-942C-4	Sequence 4, Appli
65	8	0.9	1181	4	US-09-252-991A-18714	Sequence 18714, A
66	8	0.9	1705	4	US-08-669-785-4	Sequence 4, Appli
67	7	0.8	9	1	US-08-215-805A-66	Sequence 66, Appl
68	7	0.8	23	2	US-08-295-643-11	Sequence 11, Appl
69	7	0.8	23	3	US-08-473-265-10	Sequence 10, Appl
70	7	0.8	23	3	US-08-284-747-10	Sequence 10, Appl
71	7	0.8	25	1	US-08-139-711-1	Sequence 1, Appli
72	7	0.8	44	4	US-08-469-260A-519	Sequence 519, App
73	7	0.8	44	4	US-08-488-446-519	Sequence 519, App
74	7	0.8	44	4	US-08-467-344A-519	Sequence 519, App
75	7	0.8	52	3	US-08-630-916A-78	Sequence 78, Appl
76	7	0.8	103	4	US-09-071-035-436	Sequence 436, App
77	7	0.8	103	4	US-09-252-991A-31182	Sequence 31182, A
78	7	0.8	103	4	US-09-732-210-765	Sequence 765, App
79	7	0.8	103	4	US-09-732-210-766	Sequence 766, App
80	7	0.8	104	4	US-09-732-210-761	Sequence 761, App
81	7	0.8	128	4	US-09-071-035-434	Sequence 434, App
82	7	0.8	151	4	US-09-886-319A-58	Sequence 58, Appl
83	7	0.8	153	2	US-08-387-942C-49	Sequence 49, Appl
84	7	0.8	153	4	US-09-886-319A-57	Sequence 57, Appl
85	7	0.8	210	3	US-08-856-841-9	Sequence 9, Appli
86	7	0.8	215	1	US-08-044-621D-34	Sequence 34, Appl
87	7	0.8	227	3	US-08-856-841-15	Sequence 15, Appl
88	7	0.8	233	4	US-09-252-991A-23558	Sequence 23558, A
89	7	0.8	236	4	US-09-252-991A-27285	Sequence 27285, A
90	7	0.8	242	4	US-09-252-991A-24211	Sequence 24211, A
91	7	0.8	250	3	US-08-856-841-12	Sequence 12, Appl
92	7	0.8	263	4	US-09-107-532A-6895	Sequence 6895, Ap
93	7	0.8	274	4	US-09-266-965-125	Sequence 125, App
94	7	0.8	308	4	US-09-252-991A-23093	Sequence 23093, A
95	7	0.8	315	4	US-09-134-001C-4521	Sequence 4521, Ap
96	7	0.8	327	4	US-09-252-991A-26733	Sequence 26733, A
97	7	0.8	332	4	US-09-107-532A-5477	Sequence 5477, Ap
98	7	0.8	338	4	US-09-328-352-7151	Sequence 7151, Ap
99	7	0.8	373	4	US-09-252-991A-31832	Sequence 31832, A
100	7	0.8	380	3	US-09-150-133-9	Sequence 9, Appli

101	7	0.8	380	3	US-09-150-141-9	Sequence 9, Appli
102	7	0.8	380	3	US-09-374-493-9	Sequence 9, Appli
103	7	0.8	380	3	US-09-374-824-9	Sequence 9, Appli
104	7	0.8	380	3	US-09-374-492-9	Sequence 9, Appli
105	7	0.8	387	3	US-08-856-841-17	Sequence 17, Appl
106	7	0.8	389	4	US-09-055-765-13	Sequence 13, Appl
107	7	0.8	392	4	US-09-107-532A-7217	Sequence 7217, Ap
108	7	0.8	424	4	US-09-134-001C-5009	Sequence 5009, Ap
109	7	0.8	437	4	US-09-107-532A-6709	Sequence 6709, Ap
110	7	0.8	438	4	US-09-252-991A-20127	Sequence 20127, A
111	7	0.8	441	4	US-09-107-532A-4173	Sequence 4173, Ap
112	7	0.8	446	4	US-09-252-991A-18924	Sequence 18924, A
113	7	0.8	473	4	US-09-252-991A-29636	Sequence 29636, A
114	7	0.8	477	3	US-08-843-982B-8	Sequence 8, Appli
115	7	0.8	478	3	US-09-272-114-1	Sequence 1, Appli
116	7	0.8	478	3	US-09-249-338-2	Sequence 2, Appli
117	7	0.8	480	4	US-09-252-991A-26186	Sequence 26186, A
118	7	0.8	486	3	US-08-856-841-19	Sequence 19, Appl
119	7	0.8	500	3	US-08-856-841-16	Sequence 16, Appl
120	7	0.8	500	4	US-09-252-991A-25053	Sequence 25053, A
121	7	0.8	501	3	US-08-856-841-18	Sequence 18, Appl
122	7	0.8	514	3	US-09-413-814-82	Sequence 82, Appl
123	7	0.8	514	3	US-09-413-814-94	Sequence 94, Appl
124	7	0.8	521	2	US-08-721-684C-2	Sequence 2, Appli
125	7	0.8	521	2	US-09-005-970-2	Sequence 2, Appli
126	7	0.8	521	3	US-09-407-715-2	Sequence 2, Appli
127	7	0.8	523	4	US-09-325-932A-183	Sequence 183, App
128	7	0.8	534	4	US-09-252-991A-27300	Sequence 27300, A
129	7	0.8	557	3	US-08-927-219-139	Sequence 139, App
130	7	0.8	609	3	US-08-927-219-129	Sequence 129, App
131	7	0.8	679	4	US-09-252-991A-28887	Sequence 28887, A
132	7	0.8	707	4	US-09-252-991A-24045	Sequence 24045, A
133	7	0.8	711	3	US-08-949-588-2	Sequence 2, Appli
134	7	0.8	720	3	US-09-296-284-25	Sequence 25, Appl
135	7	0.8	725	4	US-09-252-991A-27825	Sequence 27825, A
136	7	0.8	748	4	US-09-252-991A-31491	Sequence 31491, A
137	7	0.8	754	3	US-09-296-284-4	Sequence 4, Appli
138	7	0.8	783	4	US-09-252-991A-19698	Sequence 19698, A
139	7	0.8	906	4	US-09-328-352-6037	Sequence 6037, Ap
140	7	0.8	1021	4	US-09-252-991A-27405	Sequence 27405, A
141	7	0.8	1048	4	US-09-252-991A-29368	Sequence 29368, A
142	7	0.8	1075	2	US-08-993-228-19	Sequence 19, Appl
143	7	0.8	1083	3	US-08-895-601-5	Sequence 5, Appli
144	7	0.8	1138	1	US-07-973-320-2	Sequence 2, Appli
145	7	0.8	1138	1	US-07-973-320-4	Sequence 4, Appli
146	7	0.8	1242	4	US-09-107-532A-5241	Sequence 5241, Ap
147	7	0.8	1287	1	US-08-200-232-2	Sequence 2, Appli
148	7	0.8	1287	5	PCT-US95-02219-2	Sequence 2, Appli
149	7	0.8	1287	5	PCT-US95-02219A-2	Sequence 2, Appli
150	7	0.8	1296	3	US-08-470-260-3	Sequence 3, Appli
151	7	0.8	1296	3	US-08-471-491-3	Sequence 3, Appli
152	7	0.8	1296	3	US-08-466-662-3	Sequence 3, Appli
153	7	0.8	1380	4	US-09-328-352-8132	Sequence 8132, Ap
154	7	0.8	3135	1	US-08-323-170B-2	Sequence 2, Appli
155	7	0.8	3135	4	US-08-954-441-2	Sequence 2, Appli
156	6	0.6	8	3	US-09-082-279B-1438	Sequence 1438, Ap
157	6	0.6	8	4	US-09-834-784-1438	Sequence 1438, Ap
158	6	0.6	9	1	US-08-215-805A-3	Sequence 3, Appli
159	6	0.6	9	1	US-08-215-805A-6	Sequence 6, Appli
160	6	0.6	9	1	US-08-215-805A-7	Sequence 7, Appli
161	6	0.6	9	1	US-08-215-805A-9	Sequence 9, Appli
162	6	0.6	9	1	US-08-215-805A-22	Sequence 22, Appl
163	6	0.6	9	1	US-08-215-805A-24	Sequence 24, Appl
164	6	0.6	9	1	US-08-215-805A-27	Sequence 27, Appl
165	6	0.6	9	1	US-08-215-805A-32	Sequence 32, Appl
166	6	0.6	9	1	US-08-215-805A-41	Sequence 41, Appl
167	6	0.6	9	1	US-08-215-805A-62	Sequence 62, Appl
168	6	0.6	9	1	US-08-215-805A-70	Sequence 70, Appl
169	6	0.6	9	1	US-08-215-805A-75	Sequence 75, Appl
170	6	0.6	9	5	PCT-US91-03056-19	Sequence 19, Appl
171	6	0.6	10	4	US-09-517-866-9	Sequence 9, Appli
172	6	0.6	12	2	US-08-417-174-30	Sequence 30, Appl
173	6	0.6	12	2	US-08-231-565A-30	Sequence 30, Appl
174	6	0.6	12	2	US-09-007-961-30	Sequence 30, Appl
175	6	0.6	12	3	US-09-267-439-30	Sequence 30, Appl
176	6	0.6	12	4	US-09-073-138-30	Sequence 30, Appl
177	6	0.6	13	1	US-08-057-167-16	Sequence 16, Appl
178	6	0.6	13	5	PCT-US93-05412-16	Sequence 16, Appl
179	6	0.6	15	2	US-08-432-871C-81	Sequence 81, Appl
180	6	0.6	15	4	US-09-270-956-81	Sequence 81, Appl
181	6	0.6	15	4	US-09-517-866-6	Sequence 6, Appli
182	6	0.6	15	6	5304631-8	Patent No. 5304631
183	6	0.6	15	6	5304631-16	Patent No. 5304631
184	6	0.6	16	3	US-09-413-452-14	Sequence 14, Appl
185	6	0.6	16	3	US-09-413-068-14	Sequence 14, Appl
186	6	0.6	17	4	US-09-570-921-142	Sequence 142, App
187	6	0.6	18	1	US-08-086-335C-3	Sequence 3, Appli
188	6	0.6	18	3	US-09-007-905-52	Sequence 52, Appl
189	6	0.6	18	4	US-09-232-074-52	Sequence 52, Appl
190	6	0.6	20	1	US-07-777-715-4	Sequence 4, Appli
191	6	0.6	20	1	US-07-908-253-5	Sequence 5, Appli
192	6	0.6	20	1	US-08-086-335C-6	Sequence 6, Appli
193	6	0.6	20	1	US-08-170-126-6	Sequence 6, Appli
194	6	0.6	20	2	US-08-535-837-5	Sequence 5, Appli
195	6	0.6	20	3	US-08-954-418-6	Sequence 6, Appli
196	6	0.6	20	3	US-08-825-852-51	Sequence 51, Appl
197	6	0.6	20	3	US-09-007-905-59	Sequence 59, Appl
198	6	0.6	20	3	US-09-007-905-64	Sequence 64, Appl
199	6	0.6	20	3	US-09-052-888-52	Sequence 52, Appl
200	6	0.6	20	4	US-09-232-074-59	Sequence 59, Appl
201	6	0.6	20	4	US-09-232-074-64	Sequence 64, Appl
202	6	0.6	20	6	5476657-6	Patent No. 5476657
203	6	0.6	21	1	US-08-086-335C-4	Sequence 4, Appli
204	6	0.6	22	6	5489517-6	Patent No.5489517
205	6	0.6	23	1	US-08-141-324-23	Sequence 23, Appl
206	6	0.6	23	1	US-08-541-902-23	Sequence 23, Appl
207	6	0.6	23	3	US-08-448-194-16	Sequence 16, Appl
208	6	0.6	23	4	US-08-867-921-16	Sequence 16, Appl
209	6	0.6	23	4	US-09-570-921-65	Sequence 65, Appl
210	6	0.6	24	1	US-08-133-011-9	Sequence 9, Appli
211	6	0.6	24	1	US-08-322-730A-9	Sequence 9, Appli
212	6	0.6	24	1	US-08-387-874-9	Sequence 9, Appli
213	6	0.6	24	2	US-08-468-790-10	Sequence 10, Appl
214	6	0.6	24	2	US-08-383-619-9	Sequence 9, Appli
215	6	0.6	24	3	US-08-907-739-9	Sequence 9, Appli
216	6	0.6	24	4	US-09-729-597-9	Sequence 9, Appli
217	6	0.6	24	4	US-08-495-209-5	Sequence 5, Appli
218	6	0.6	24	5	PCT-US93-08364-9	Sequence 9, Appli
219	6	0.6	24	5	PCT-US96-10905-5	Sequence 5, Appli
220	6	0.6	25	1	US-08-475-989-9	Sequence 9, Appli
221	6	0.6	25	2	US-08-475-985-9	Sequence 9, Appli
222	6	0.6	25	2	US-08-472-172-21	Sequence 21, Appl
223	6	0.6	25	3	US-08-256-839-9	Sequence 9, Appli
224	6	0.6	26	1	US-07-971-096-12	Sequence 12, Appl
225	6	0.6	26	1	US-08-175-096-12	Sequence 12, Appl
226	6	0.6	26	4	US-09-461-325-383	Sequence 383, App
227	6	0.6	28	3	US-09-413-452-15	Sequence 15, Appl
228	6	0.6	28	3	US-09-413-068-15	Sequence 15, Appl
229	6	0.6	30	3	US-08-448-489-7	Sequence 7, Appli
230	6	0.6	32	6	5464756-12	Patent No. 5464756
231	6	0.6	33	3	US-08-810-712-25	Sequence 25, Appl
232	6	0.6	33	4	US-09-181-941-14	Sequence 14, Appl
233	6	0.6	34	2	US-08-248-839C-165	Sequence 165, App
234	6	0.6	38	3	US-09-007-905-48	Sequence 48, Appl
235	6	0.6	38	4	US-09-232-074-48	Sequence 48, Appl
236	6	0.6	40	2	US-08-637-759B-484	Sequence 484, App
237	6	0.6	40	3	US-08-871-355A-484	Sequence 484, App
238	6	0.6	40	4	US-09-201-945-484	Sequence 484, App
239	6	0.6	43	3	US-09-007-905-57	Sequence 57, Appl
240	6	0.6	43	4	US-09-232-074-57	Sequence 57, Appl
241	6	0.6	45	3	US-09-100-600A-87	Sequence 87, Appl
242	6	0.6	45	3	US-09-100-600A-90	Sequence 90, Appl
243	6	0.6	46	3	US-08-857-076-83	Sequence 83, Appl
244	6	0.6	48	1	US-08-062-472B-46	Sequence 46, Appl
245	6	0.6	48	4	US-09-205-258-358	Sequence 358, App
246	6	0.6	50	4	US-09-447-125B-19	Sequence 19, Appl

247	6	0.6	50	5	PCT-US93-05640-33	Sequence 33, Appl
248	6	0.6	51	1	US-08-346-849-14	Sequence 14, Appl
249	6	0.6	51	2	US-08-293-284A-14	Sequence 14, Appl
250	6	0.6	51	4	US-09-187-789-34	Sequence 34, Appl
251	6	0.6	51	4	US-09-139-600-29	Sequence 29, Appl
252	6	0.6	51	4	US-08-898-300-14	Sequence 14, Appl
253	6	0.6	52	1	US-08-247-475-39	Sequence 39, Appl
254	6	0.6	52	1	US-08-479-650-39	Sequence 39, Appl
255	6	0.6	52	1	US-08-191-866D-61	Sequence 61, Appl
256	6	0.6	52	1	US-08-674-169-39	Sequence 39, Appl
257	6	0.6	52	2	US-08-185-949B-61	Sequence 61, Appl
258	6	0.6	58	4	US-08-963-851-22	Sequence 22, Appl
259	6	0.6	59	3	US-08-812-586-7	Sequence 7, Appli
260	6	0.6	59	4	US-09-535-832A-7	Sequence 7, Appli
261	6	0.6	60	2	US-08-417-210A-113	Sequence 113, App
262	6	0.6	60	3	US-08-812-586-6	Sequence 6, Appli
263	6	0.6	60	4	US-09-535-832A-6	Sequence 6, Appli
264	6	0.6	61	4	US-09-134-001C-4272	Sequence 4272, Ap
265	6	0.6	63	3	US-08-817-787-17	Sequence 17, Appl
266	6	0.6	63	6	5198542-11	Patent No. 5198542
267	6	0.6	64	4	US-09-205-258-570	Sequence 570, App
268	6	0.6	65	3	US-09-007-905-36	Sequence 36, Appl
269	6	0.6	65	4	US-09-232-074-36	Sequence 36, Appl
270	6	0.6	66	3	US-09-007-905-40	Sequence 40, Appl
271	6	0.6	66	4	US-09-232-074-40	Sequence 40, Appl
272	6	0.6	69	3	US-09-007-905-30	Sequence 30, Appl
273	6	0.6	69	4	US-09-232-074-30	Sequence 30, Appl
274	6	0.6	75	4	US-09-252-991A-28889	Sequence 28889, A
275	6	0.6	75	6	5252466-3	Patent No. 5252466
276	6	0.6	77	4	US-08-936-165A-511	Sequence 511, App
277	6	0.6	81	4	US-09-198-452A-1167	Sequence 1167, Ap
278	6	0.6	81	4	US-09-107-532A-4732	Sequence 4732, Ap
279	6	0.6	82	4	US-08-858-207A-272	Sequence 272, App
280	6	0.6	83	2	US-08-858-767-33	Sequence 33, Appl
281	6	0.6	83	2	US-08-858-767-34	Sequence 34, Appl
282	6	0.6	83	2	US-08-863-028-33	Sequence 33, Appl
283	6	0.6	83	2	US-08-863-028-34	Sequence 34, Appl
284	6	0.6	87	4	US-09-191-468-45	Sequence 45, Appl
285	6	0.6	87	4	US-09-191-468-47	Sequence 47, Appl
286	6	0.6	87	4	US-09-191-468-49	Sequence 49, Appl
287	6	0.6	87	4	US-09-191-468-51	Sequence 51, Appl
288	6	0.6	87	4	US-09-191-468-53	Sequence 53, Appl
289	6	0.6	90	4	US-09-328-352-6160	Sequence 6160, Ap
290	6	0.6	91	3	US-08-905-223-409	Sequence 409, App
291	6	0.6	94	1	US-08-266-451B-20	Sequence 20, Appl
292	6	0.6	94	2	US-08-748-725-20	Sequence 20, Appl
293	6	0.6	94	2	US-09-047-125-32	Sequence 32, Appl
294	6	0.6	94	3	US-07-736-335E-32	Sequence 32, Appl
295	6	0.6	94	4	US-09-461-325-381	Sequence 381, App
296	6	0.6	95	1	US-08-456-265A-107	Sequence 107, App
297	6	0.6	95	1	US-08-456-265A-109	Sequence 109, App
298	6	0.6	95	1	US-08-456-265A-110	Sequence 110, App
299	6	0.6	95	2	US-08-971-217-107	Sequence 107, App
300	6	0.6	95	2	US-08-971-217-109	Sequence 109, App
301	6	0.6	95	2	US-08-971-217-110	Sequence 110, App
302	6	0.6	95	3	US-09-350-600-107	Sequence 107, App
303	6	0.6	95	3	US-09-350-600-109	Sequence 109, App
304	6	0.6	95	3	US-09-350-600-110	Sequence 110, App
305	6	0.6	98	4	US-09-107-532A-7167	Sequence 7167, Ap
306	6	0.6	99	2	US-08-537-400-19	Sequence 19, Appl
307	6	0.6	99	4	US-09-370-838-95	Sequence 95, Appl
308	6	0.6	100	2	US-08-464-517-8	Sequence 8, Appli
309	6	0.6	100	2	US-08-246-361A-8	Sequence 8, Appli
310	6	0.6	100	3	US-08-463-772-8	Sequence 8, Appli
311	6	0.6	100	5	PCT-US93-05000-8	Sequence 8, Appli
312	6	0.6	101	6	5164490-5	Patent No. 5164490
313	6	0.6	103	4	US-09-732-210-763	Sequence 763, App
314	6	0.6	103	4	US-09-732-210-1231	Sequence 1231, Ap
315	6	0.6	104	3	US-08-857-076-112	Sequence 112, App
316	6	0.6	105	3	US-08-867-381A-4	Sequence 4, Appli
317	6	0.6	105	4	US-09-521-144-4	Sequence 4, Appli
318	6	0.6	105	4	US-09-732-210-781	Sequence 781, App
319	6	0.6	106	2	US-08-464-517-25	Sequence 25, Appl
320	6	0.6	106	2	US-08-246-361A-25	Sequence 25, Appl
321	6	0.6	106	3	US-08-463-772-25	Sequence 25, Appl
322	6	0.6	106	5	PCT-US93-05000-25	Sequence 25, Appl
323	6	0.6	107	4	US-09-134-001C-4474	Sequence 4474, Ap
324	6	0.6	107	4	US-09-186-276B-63	Sequence 63, Appl
325	6	0.6	107	4	US-08-842-445-63	Sequence 63, Appl
326	6	0.6	107	4	US-09-186-188B-63	Sequence 63, Appl
327	6	0.6	111	1	US-08-456-265A-111	Sequence 111, App
328	6	0.6	111	2	US-08-971-217-111	Sequence 111, App
329	6	0.6	111	3	US-09-350-600-111	Sequence 111, App
330	6	0.6	111	4	US-08-469-260A-72	Sequence 72, Appl
331	6	0.6	111	4	US-08-488-446-72	Sequence 72, Appl
332	6	0.6	111	4	US-08-467-344A-72	Sequence 72, Appl
333	6	0.6	112	4	US-09-252-991A-22828	Sequence 22828, A
334	6	0.6	115	1	US-08-209-747-17	Sequence 17, Appl
335	6	0.6	115	1	US-08-458-298-17	Sequence 17, Appl
336	6	0.6	116	1	US-08-209-747-18	Sequence 18, Appl
337	6	0.6	116	1	US-08-458-298-18	Sequence 18, Appl
338	6	0.6	116	4	US-09-252-991A-23793	Sequence 23793, A
339	6	0.6	116	4	US-09-732-210-661	Sequence 661, App
340	6	0.6	118	4	US-09-732-210-158	Sequence 158, App
341	6	0.6	119	4	US-09-328-352-5881	Sequence 5881, Ap
342	6	0.6	120	2	US-08-637-759B-269	Sequence 269, App
343	6	0.6	120	3	US-08-871-355A-269	Sequence 269, App
344	6	0.6	120	4	US-09-201-945-269	Sequence 269, App
345	6	0.6	122	4	US-09-462-606-65	Sequence 65, Appl
346	6	0.6	123	3	US-08-840-316-3	Sequence 3, Appli
347	6	0.6	123	3	US-08-478-507-9	Sequence 9, Appli
348	6	0.6	123	3	US-08-809-523-3	Sequence 3, Appli
349	6	0.6	123	3	US-09-128-275A-9	Sequence 9, Appli
350	6	0.6	123	3	US-08-471-971-3	Sequence 3, Appli
351	6	0.6	123	4	US-09-553-427-9	Sequence 9, Appli
352	6	0.6	123	4	US-09-462-606-13	Sequence 13, Appl
353	6	0.6	123	4	US-09-462-606-58	Sequence 58, Appl
354	6	0.6	123	4	US-09-462-606-59	Sequence 59, Appl
355	6	0.6	123	4	US-09-462-606-60	Sequence 60, Appl
356	6	0.6	123	4	US-09-462-606-61	Sequence 61, Appl
357	6	0.6	123	4	US-09-462-606-62	Sequence 62, Appl
358	6	0.6	123	4	US-09-462-606-63	Sequence 63, Appl
359	6	0.6	123	4	US-09-462-606-64	Sequence 64, Appl
360	6	0.6	123	4	US-08-469-260A-61	Sequence 61, Appl
361	6	0.6	123	4	US-09-402-776-3	Sequence 3, Appli
362	6	0.6	123	4	US-09-172-699-4	Sequence 4, Appli
363	6	0.6	123	4	US-08-488-446-61	Sequence 61, Appl
364	6	0.6	123	4	US-08-467-344A-61	Sequence 61, Appl
365	6	0.6	123	5	PCT-US93-08849A-3	Sequence 3, Appli
366	6	0.6	123	5	PCT-US93-08849-3	Sequence 3, Appli
367	6	0.6	124	1	US-08-240-049B-19	Sequence 19, Appl
368	6	0.6	124	1	US-08-240-049B-20	Sequence 20, Appl
369	6	0.6	124	3	US-08-542-634-21	Sequence 21, Appl
370	6	0.6	124	3	US-08-542-634-22	Sequence 22, Appl
371	6	0.6	124	3	US-08-477-292-21	Sequence 21, Appl
372	6	0.6	124	3	US-08-477-292-22	Sequence 22, Appl
373	6	0.6	124	4	US-09-252-991A-31009	Sequence 31009, A
374	6	0.6	124	5	PCT-US95-13703-21	Sequence 21, Appl
375	6	0.6	124	5	PCT-US95-13703-22	Sequence 22, Appl
376	6	0.6	127	3	US-08-806-121B-3	Sequence 3, Appli
377	6	0.6	127	3	US-08-812-586-47	Sequence 47, Appl
378	6	0.6	127	4	US-09-443-061-3	Sequence 3, Appli
379	6	0.6	127	4	US-09-535-832A-44	Sequence 44, Appl
380	6	0.6	127	4	US-09-732-210-1685	Sequence 1685, Ap
381	6	0.6	127	4	US-09-107-532A-6784	Sequence 6784, Ap
382	6	0.6	129	3	US-09-476-482-8	Sequence 8, Appli
383	6	0.6	130	2	US-08-771-602D-44	Sequence 44, Appl
384	6	0.6	132	1	US-07-820-154A-2	Sequence 2, Appli
385	6	0.6	132	1	US-07-820-154A-6	Sequence 6, Appli
386	6	0.6	132	2	US-08-097-554A-2	Sequence 2, Appli
387	6	0.6	132	2	US-08-097-554A-6	Sequence 6, Appli
388	6	0.6	132	3	US-08-480-640A-2	Sequence 2, Appli
389	6	0.6	132	3	US-08-480-640A-6	Sequence 6, Appli
390	6	0.6	132	3	US-08-295-802-2	Sequence 2, Appli
391	6	0.6	132	3	US-08-295-802-6	Sequence 6, Appli
392	6	0.6	132	3	US-08-686-968C-102	Sequence 102, App

393	6	0.6	132	3	US-08-488-237A-2	Sequence 2, Appli
394	6	0.6	132	3	US-08-488-237A-6	Sequence 6, Appli
395	6	0.6	132	4	US-08-375-992A-2	Sequence 2, Appli
396	6	0.6	132	4	US-08-375-992A-6	Sequence 6, Appli
397	6	0.6	132	4	US-08-472-679H-2	Sequence 2, Appli
398	6	0.6	132	4	US-08-472-679H-6	Sequence 6, Appli
399	6	0.6	132	5	PCT-US93-00324-2	Sequence 2, Appli
400	6	0.6	132	5	PCT-US93-00324-6	Sequence 6, Appli
401	6	0.6	133	1	US-07-800-366-1	Sequence 1, Appli
402	6	0.6	133	1	US-08-354-456A-5	Sequence 5, Appli
403	6	0.6	133	1	US-08-354-456A-6	Sequence 6, Appli
404	6	0.6	133	1	US-08-225-224-3	Sequence 3, Appli
405	6	0.6	133	1	US-08-318-193-89	Sequence 89, Appl
406	6	0.6	133	1	US-08-284-393B-1	Sequence 1, Appli
407	6	0.6	133	1	US-08-284-393B-2	Sequence 2, Appli
408	6	0.6	133	1	US-08-284-393B-3	Sequence 3, Appli
409	6	0.6	133	1	US-08-734-471-1	Sequence 1, Appli
410	6	0.6	133	3	US-08-722-258-3	Sequence 3, Appli
411	6	0.6	133	3	US-08-905-223-412	Sequence 412, App
412	6	0.6	133	3	US-08-817-787-13	Sequence 13, Appl
413	6	0.6	133	4	US-09-570-921-54	Sequence 54, Appl
414	6	0.6	133	4	US-09-252-991A-29319	Sequence 29319, A
415	6	0.6	133	4	US-09-538-873-2	Sequence 2, Appli
416	6	0.6	133	5	PCT-US95-04468-3	Sequence 3, Appli
417	6	0.6	133	5	PCT-US95-08950-1	Sequence 1, Appli
418	6	0.6	133	5	PCT-US95-08950-2	Sequence 2, Appli
419	6	0.6	133	5	PCT-US95-08950-3	Sequence 3, Appli
420	6	0.6	133	6	5210029-1	Patent No. 5210029
421	6	0.6	133	6	5256769-1	Patent No. 5256769
422	6	0.6	133	6	5464939-2	Patent No. 5464939
423	6	0.6	134	6	5496924-55	Patent No. 5496924
424	6	0.6	135	4	US-08-858-207A-489	Sequence 489, App
425	6	0.6	136	2	US-08-580-545B-8	Sequence 8, Appli
426	6	0.6	136	3	US-09-262-653A-8	Sequence 8, Appli
427	6	0.6	138	4	US-09-328-352-7972	Sequence 7972, Ap
428	6	0.6	141	1	US-07-695-564-5	Sequence 5, Appli
429	6	0.6	141	1	US-08-241-387-5	Sequence 5, Appli
430	6	0.6	141	2	US-08-690-011A-43	Sequence 43, Appl
431	6	0.6	141	4	US-09-299-495F-43	Sequence 43, Appl
432	6	0.6	141	4	US-09-732-210-758	Sequence 758, App
433	6	0.6	143	4	US-09-252-991A-26817	Sequence 26817, A
434	6	0.6	145	2	US-08-686-599A-20	Sequence 20, Appl
435	6	0.6	145	4	US-09-134-001C-5243	Sequence 5243, Ap
436	6	0.6	145	4	US-09-252-991A-22315	Sequence 22315, A
437	6	0.6	146	4	US-09-252-991A-25506	Sequence 25506, A
438	6	0.6	147	4	US-09-252-991A-17088	Sequence 17088, A
439	6	0.6	148	1	US-08-207-904-15	Sequence 15, Appl
440	6	0.6	149	1	US-07-695-564-7	Sequence 7, Appli
441	6	0.6	149	1	US-08-241-387-7	Sequence 7, Appli
442	6	0.6	149	4	US-09-252-991A-30582	Sequence 30582, A
443	6	0.6	150	2	US-08-460-694-3	Sequence 3, Appli
444	6	0.6	150	2	US-08-387-942C-44	Sequence 44, Appl
445	6	0.6	150	2	US-08-387-942C-52	Sequence 52, Appl
446	6	0.6	150	3	US-08-460-744-3	Sequence 3, Appli
447	6	0.6	150	3	US-07-667-711B-3	Sequence 3, Appli
448	6	0.6	150	3	US-09-211-542A-14	Sequence 14, Appl
449	6	0.6	151	2	US-08-387-942C-42	Sequence 42, Appl
450	6	0.6	152	4	US-09-107-532A-3744	Sequence 3744, Ap
451	6	0.6	153	2	US-08-387-942C-10	Sequence 10, Appl
452	6	0.6	153	2	US-08-387-942C-43	Sequence 43, Appl
453	6	0.6	153	3	US-09-012-366-3	Sequence 3, Appli
454	6	0.6	153	3	US-08-759-628-8	Sequence 8, Appli
455	6	0.6	153	4	US-09-522-217-111	Sequence 111, App
456	6	0.6	153	4	US-09-328-352-4423	Sequence 4423, Ap
457	6	0.6	153	6	5314995-7	Patent No. 5314995
458	6	0.6	154	4	US-09-252-991A-29366	Sequence 29366, A
459	6	0.6	154	4	US-09-107-532A-4031	Sequence 4031, Ap
460	6	0.6	155	1	US-08-150-203A-8	Sequence 8, Appli
461	6	0.6	155	1	US-08-454-730-8	Sequence 8, Appli
462	6	0.6	155	4	US-08-949-788-8	Sequence 8, Appli
463	6	0.6	155	4	US-09-252-991A-26113	Sequence 26113, A
464	6	0.6	156	4	US-09-134-001C-5252	Sequence 5252, Ap
465	6	0.6	157	1	US-08-450-065-2	Sequence 2, Appli
466	6	0.6	157	1	US-08-450-595-2	Sequence 2, Appli
467	6	0.6	157	3	US-08-818-562-2	Sequence 2, Appli
468	6	0.6	157	4	US-09-134-001C-2991	Sequence 2991, Ap
469	6	0.6	157	4	US-09-628-445-2	Sequence 2, Appli
470	6	0.6	158	2	US-08-888-497-22	Sequence 22, Appl
471	6	0.6	158	4	US-09-362-230-22	Sequence 22, Appl
472	6	0.6	158	4	US-09-252-991A-17285	Sequence 17285, A
473	6	0.6	158	4	US-09-252-991A-28490	Sequence 28490, A
474	6	0.6	158	5	PCT-US94-07926-22	Sequence 22, Appl
475	6	0.6	159	4	US-09-732-210-727	Sequence 727, App
476	6	0.6	160	4	US-09-134-001C-5659	Sequence 5659, Ap
477	6	0.6	162	2	US-08-716-317-10	Sequence 10, Appl
478	6	0.6	163	2	US-08-716-317-15	Sequence 15, Appl
479	6	0.6	163	2	US-08-846-762-75	Sequence 75, Appl
480	6	0.6	163	3	US-08-990-791-11	Sequence 11, Appl
481	6	0.6	163	4	US-09-724-623-107	Sequence 107, App
482	6	0.6	163	4	US-09-372-591-11	Sequence 11, Appl
483	6	0.6	164	1	US-08-318-193-60	Sequence 60, Appl
484	6	0.6	164	4	US-09-252-991A-16845	Sequence 16845, A
485	6	0.6	166	4	US-09-499-148-1	Sequence 1, Appli
486	6	0.6	166	4	US-09-499-148-4	Sequence 4, Appli
487	6	0.6	167	1	US-08-246-427A-2	Sequence 2, Appli
488	6	0.6	167	2	US-08-766-620-2	Sequence 2, Appli
489	6	0.6	167	3	US-08-959-004-7	Sequence 7, Appli
490	6	0.6	167	5	PCT-US95-06094-2	Sequence 2, Appli
491	6	0.6	168	2	US-08-353-476-73	Sequence 73, Appl
492	6	0.6	168	4	US-09-252-991A-27144	Sequence 27144, A
493	6	0.6	168	4	US-09-252-991A-32471	Sequence 32471, A
494	6	0.6	168	4	US-09-328-352-7984	Sequence 7984, Ap
495	6	0.6	171	4	US-09-134-001C-3256	Sequence 3256, Ap
496	6	0.6	171	4	US-09-328-352-4243	Sequence 4243, Ap
497	6	0.6	172	3	US-08-149-101A-21	Sequence 21, Appl
498	6	0.6	172	5	PCT-US94-12873-21	Sequence 21, Appl
499	6	0.6	173	1	US-08-193-977-4	Sequence 4, Appli
500	6	0.6	173	4	US-09-107-532A-5804	Sequence 5804, Ap
501	6	0.6	174	3	US-08-149-101A-20	Sequence 20, Appl
502	6	0.6	174	4	US-09-252-991A-21312	Sequence 21312, A
503	6	0.6	174	5	PCT-US94-12873-20	Sequence 20, Appl
504	6	0.6	176	1	US-07-924-054-8	Sequence 8, Appli
505	6	0.6	176	1	US-08-076-011-2	Sequence 2, Appli
506	6	0.6	176	1	US-08-062-472B-20	Sequence 20, Appl
507	6	0.6	176	3	US-09-232-446B-4	Sequence 4, Appli
508	6	0.6	176	6	5198542-4	Patent No. 5198542
509	6	0.6	178	3	US-09-232-446B-5	Sequence 5, Appli
510	6	0.6	179	1	US-08-193-977-6	Sequence 6, Appli
511	6	0.6	179	3	US-08-867-381A-52	Sequence 52, Appl
512	6	0.6	179	4	US-09-521-144-52	Sequence 52, Appl
513	6	0.6	181	1	US-08-209-182C-4	Sequence 4, Appli
514	6	0.6	181	3	US-08-961-083-114	Sequence 114, App
515	6	0.6	181	4	US-09-536-784-114	Sequence 114, App
516	6	0.6	181	5	PCT-US92-05612-4	Sequence 4, Appli
517	6	0.6	182	3	US-08-149-101A-25	Sequence 25, Appl
518	6	0.6	182	5	PCT-US94-12873-25	Sequence 25, Appl
519	6	0.6	183	1	US-08-009-973-1	Sequence 1, Appli
520	6	0.6	184	1	US-08-567-047-2	Sequence 2, Appli
521	6	0.6	184	2	US-08-693-182-2	Sequence 2, Appli
522	6	0.6	184	2	US-08-567-048-2	Sequence 2, Appli
523	6	0.6	184	2	US-09-008-482-2	Sequence 2, Appli
524	6	0.6	184	2	US-08-945-529-8	Sequence 8, Appli
525	6	0.6	184	2	US-08-945-529-9	Sequence 9, Appli
526	6	0.6	184	2	US-08-945-529-10	Sequence 10, Appl
527	6	0.6	184	2	US-08-945-529-11	Sequence 11, Appl
528	6	0.6	184	2	US-08-945-529-12	Sequence 12, Appl
529	6	0.6	184	3	US-08-149-101A-17	Sequence 17, Appl
530	6	0.6	184	3	US-08-149-101A-22	Sequence 22, Appl
531	6	0.6	184	4	US-09-252-991A-26628	Sequence 26628, A
532	6	0.6	184	5	PCT-US94-12873-17	Sequence 17, Appl
533	6	0.6	184	5	PCT-US94-12873-22	Sequence 22, Appl
534	6	0.6	184	6	5186931-1	Patent No. 5186931
535	6	0.6	185	1	US-07-632-070B-1	Sequence 1, Appli
536	6	0.6	185	1	US-07-918-181A-2	Sequence 2, Appli
537	6	0.6	185	1	US-07-918-181A-4	Sequence 4, Appli
538	6	0.6	185	1	US-07-918-181A-6	Sequence 6, Appli

539	6	0.6	185	1	US-07-918-181A-8	Sequence 8, Appli
540	6	0.6	185	1	US-08-231-575-2	Sequence 2, Appli
541	6	0.6	185	1	US-08-231-575-4	Sequence 4, Appli
542	6	0.6	185	1	US-08-231-575-6	Sequence 6, Appli
543	6	0.6	185	1	US-08-231-575-8	Sequence 8, Appli
544	6	0.6	185	1	US-08-246-427A-5	Sequence 5, Appli
545	6	0.6	185	2	US-08-716-317-7	Sequence 7, Appli
546	6	0.6	185	2	US-08-766-620-5	Sequence 5, Appli
547	6	0.6	185	3	US-08-705-875A-5	Sequence 5, Appli
548	6	0.6	185	3	US-09-122-443-16	Sequence 16, Appl
549	6	0.6	185	3	US-09-220-731-22	Sequence 22, Appl
550	6	0.6	185	4	US-09-558-089-16	Sequence 16, Appl
551	6	0.6	185	4	US-09-558-087-16	Sequence 16, Appl
552	6	0.6	185	4	US-09-242-999-5	Sequence 5, Appli
553	6	0.6	185	4	US-09-198-452A-106	Sequence 106, App
554	6	0.6	185	5	PCT-US92-05612-2	Sequence 2, Appli
555	6	0.6	185	5	PCT-US93-06928-2	Sequence 2, Appli
556	6	0.6	185	5	PCT-US93-06928-4	Sequence 4, Appli
557	6	0.6	185	5	PCT-US93-06928-6	Sequence 6, Appli
558	6	0.6	185	5	PCT-US93-06928-8	Sequence 8, Appli
559	6	0.6	185	6	5186931-2	Patent No. 5186931
560	6	0.6	186	1	US-07-632-070B-2	Sequence 2, Appli
561	6	0.6	186	1	US-07-745-382-20	Sequence 20, Appl
562	6	0.6	186	1	US-07-921-848-20	Sequence 20, Appl
563	6	0.6	186	1	US-08-209-182C-2	Sequence 2, Appli
564	6	0.6	186	1	US-08-209-182C-6	Sequence 6, Appli
565	6	0.6	186	1	US-08-165-301A-20	Sequence 20, Appl
566	6	0.6	186	3	US-08-469-318-163	Sequence 163, App
567	6	0.6	186	3	US-08-468-609A-163	Sequence 163, App
568	6	0.6	186	3	US-08-810-436-20	Sequence 20, Appl
569	6	0.6	186	4	US-08-446-872A-163	Sequence 163, App
570	6	0.6	186	4	US-08-762-227A-163	Sequence 163, App
571	6	0.6	186	5	PCT-US92-05612-6	Sequence 6, Appli
572	6	0.6	186	5	PCT-US94-14179-20	Sequence 20, Appl
573	6	0.6	186	5	PCT-US95-01185-163	Sequence 163, App
574	6	0.6	187	1	US-07-632-070B-3	Sequence 3, Appli
575	6	0.6	188	3	US-09-122-443-12	Sequence 12, Appl
576	6	0.6	188	4	US-09-558-089-12	Sequence 12, Appl
577	6	0.6	188	4	US-09-558-087-12	Sequence 12, Appl
578	6	0.6	189	1	US-08-152-485-4	Sequence 4, Appli
579	6	0.6	189	1	US-08-463-089-4	Sequence 4, Appli
580	6	0.6	189	1	US-08-461-360A-4	Sequence 4, Appli
581	6	0.6	189	1	US-08-461-359-4	Sequence 4, Appli
582	6	0.6	189	4	US-08-671-548C-52	Sequence 52, Appl
583	6	0.6	189	5	PCT-US94-12904-4	Sequence 4, Appli
584	6	0.6	191	3	US-08-764-563-1	Sequence 1, Appli
585	6	0.6	191	3	US-08-720-625-2	Sequence 2, Appli
586	6	0.6	191	4	US-09-183-861-82	Sequence 82, Appl
587	6	0.6	191	4	US-09-022-765-82	Sequence 82, Appl
588	6	0.6	191	4	US-09-551-974A-82	Sequence 82, Appl
589	6	0.6	192	1	US-08-086-428B-81	Sequence 81, Appl
590	6	0.6	192	1	US-08-086-428B-82	Sequence 82, Appl
591	6	0.6	192	1	US-08-086-428B-83	Sequence 83, Appl
592	6	0.6	192	1	US-08-086-428B-84	Sequence 84, Appl
593	6	0.6	192	2	US-08-468-570-81	Sequence 81, Appl
594	6	0.6	192	2	US-08-468-570-82	Sequence 82, Appl
595	6	0.6	192	2	US-08-468-570-83	Sequence 83, Appl
596	6	0.6	192	2	US-08-468-570-84	Sequence 84, Appl
597	6	0.6	192	2	US-08-290-665A-81	Sequence 81, Appl
598	6	0.6	192	2	US-08-290-665A-82	Sequence 82, Appl
599	6	0.6	192	2	US-08-290-665A-83	Sequence 83, Appl
600	6	0.6	192	2	US-08-290-665A-84	Sequence 84, Appl
601	6	0.6	192	3	US-09-370-807-10	Sequence 10, Appl
602	6	0.6	192	4	US-09-921-259-10	Sequence 10, Appl
603	6	0.6	192	4	US-08-466-601A-81	Sequence 81, Appl
604	6	0.6	192	4	US-08-466-601A-82	Sequence 82, Appl
605	6	0.6	192	4	US-08-466-601A-83	Sequence 83, Appl
606	6	0.6	192	4	US-08-466-601A-84	Sequence 84, Appl
607	6	0.6	192	5	PCT-US95-10398-81	Sequence 81, Appl
608	6	0.6	192	5	PCT-US95-10398-82	Sequence 82, Appl
609	6	0.6	192	5	PCT-US95-10398-83	Sequence 83, Appl
610	6	0.6	192	5	PCT-US95-10398-84	Sequence 84, Appl
611	6	0.6	193	1	US-08-248-466B-14	Sequence 14, Appl
612	6	0.6	193	4	US-09-328-352-6034	Sequence 6034, Ap
613	6	0.6	194	1	US-08-063-552-8	Sequence 8, Appli
614	6	0.6	194	5	PCT-US93-05704-8	Sequence 8, Appli
615	6	0.6	195	4	US-09-252-991A-18622	Sequence 18622, A
616	6	0.6	196	4	US-09-252-991A-19503	Sequence 19503, A
617	6	0.6	196	4	US-09-252-991A-26430	Sequence 26430, A
618	6	0.6	197	4	US-09-125-619-25	Sequence 25, Appl
619	6	0.6	198	4	US-09-252-991A-23791	Sequence 23791, A
620	6	0.6	199	4	US-09-252-991A-31679	Sequence 31679, A
621	6	0.6	199	4	US-09-107-532A-6773	Sequence 6773, Ap
622	6	0.6	200	1	US-07-855-412B-1	Sequence 1, Appli
623	6	0.6	200	1	US-08-209-182C-8	Sequence 8, Appli
624	6	0.6	200	2	US-08-308-887A-1	Sequence 1, Appli
625	6	0.6	200	3	US-08-881-094-1	Sequence 1, Appli
626	6	0.6	200	4	US-09-252-991A-19724	Sequence 19724, A
627	6	0.6	200	4	US-09-252-991A-27855	Sequence 27855, A
628	6	0.6	201	4	US-09-252-991A-21340	Sequence 21340, A
629	6	0.6	201	5	PCT-US92-05612-8	Sequence 8, Appli
630	6	0.6	203	6	5168051-12	Patent No. 5168051
631	6	0.6	204	3	US-09-230-637-25	Sequence 25, Appl
632	6	0.6	204	4	US-09-230-371A-26	Sequence 26, Appl
633	6	0.6	205	4	US-09-107-532A-6070	Sequence 6070, Ap
634	6	0.6	206	4	US-09-252-991A-18746	Sequence 18746, A
635	6	0.6	207	4	US-09-199-637A-211	Sequence 211, App
636	6	0.6	208	3	US-08-630-172-15	Sequence 15, Appl
637	6	0.6	208	3	US-09-375-419-15	Sequence 15, Appl
638	6	0.6	208	4	US-09-460-384-36	Sequence 36, Appl
639	6	0.6	208	4	US-09-710-092-8	Sequence 8, Appli
640	6	0.6	209	4	US-09-252-991A-18156	Sequence 18156, A
641	6	0.6	209	4	US-09-328-352-8077	Sequence 8077, Ap
642	6	0.6	210	4	US-09-897-537A-2	Sequence 2, Appli
643	6	0.6	212	1	US-08-792-019B-9	Sequence 9, Appli
644	6	0.6	212	3	US-08-988-819-9	Sequence 9, Appli
645	6	0.6	212	3	US-09-016-534-9	Sequence 9, Appli
646	6	0.6	212	3	US-08-097-869-7	Sequence 7, Appli
647	6	0.6	212	3	US-08-795-473B-6	Sequence 6, Appli
648	6	0.6	212	3	US-09-230-637-45	Sequence 45, Appl
649	6	0.6	212	3	US-09-155-941-1	Sequence 1, Appli
650	6	0.6	212	4	US-09-230-371A-27	Sequence 27, Appl
651	6	0.6	212	4	US-09-439-856-6	Sequence 6, Appli
652	6	0.6	212	4	US-09-487-792-14	Sequence 14, Appl
653	6	0.6	212	4	US-09-908-594-14	Sequence 14, Appl
654	6	0.6	212	6	5510472-2	Patent No. 5510472
655	6	0.6	213	4	US-09-252-991A-22585	Sequence 22585, A
656	6	0.6	215	4	US-09-134-001C-4736	Sequence 4736, Ap
657	6	0.6	215	4	US-09-252-991A-27377	Sequence 27377, A
658	6	0.6	215	4	US-09-252-991A-31700	Sequence 31700, A
659	6	0.6	219	1	US-08-152-485-2	Sequence 2, Appli
660	6	0.6	219	1	US-08-463-089-2	Sequence 2, Appli
661	6	0.6	219	1	US-08-461-360A-2	Sequence 2, Appli
662	6	0.6	219	1	US-08-461-359-2	Sequence 2, Appli
663	6	0.6	219	5	PCT-US94-12904-2	Sequence 2, Appli
664	6	0.6	221	4	US-09-328-352-5680	Sequence 5680, Ap
665	6	0.6	221	4	US-09-107-532A-4477	Sequence 4477, Ap
666	6	0.6	221	4	US-09-107-532A-5877	Sequence 5877, Ap
667	6	0.6	222	4	US-09-107-532A-4165	Sequence 4165, Ap
668	6	0.6	223	2	US-08-928-612-3	Sequence 3, Appli
669	6	0.6	224	4	US-09-082-920-2	Sequence 2, Appli
670	6	0.6	225	1	US-08-290-979A-8	Sequence 8, Appli
671	6	0.6	226	4	US-09-134-001C-5494	Sequence 5494, Ap
672	6	0.6	226	4	US-09-489-847-215	Sequence 215, App
673	6	0.6	227	4	US-09-252-991A-30173	Sequence 30173, A
674	6	0.6	227	4	US-09-198-452A-1245	Sequence 1245, Ap
675	6	0.6	229	4	US-08-936-165A-523	Sequence 523, App
676	6	0.6	229	4	US-09-107-532A-4842	Sequence 4842, Ap
677	6	0.6	233	4	US-09-252-991A-25399	Sequence 25399, A
678	6	0.6	234	4	US-09-252-991A-21695	Sequence 21695, A
679	6	0.6	235	1	US-08-015-985-13	Sequence 13, Appl
680	6	0.6	235	2	US-08-378-939-12	Sequence 12, Appl
681	6	0.6	235	4	US-09-198-452A-752	Sequence 752, App
682	6	0.6	236	3	US-08-937-271-11	Sequence 11, Appl
683	6	0.6	237	2	US-08-578-709-11	Sequence 11, Appl
684	6	0.6	237	4	US-09-328-352-8181	Sequence 8181, Ap

685	6	0.6	239	4	US-09-372-422A-42	Sequence 42, Appl
686	6	0.6	239	4	US-09-252-991A-19895	Sequence 19895, A
687	6	0.6	239	4	US-09-252-991A-25446	Sequence 25446, A
688	6	0.6	240	4	US-09-489-847-143	Sequence 143, App
689	6	0.6	240	4	US-09-252-991A-26041	Sequence 26041, A
690	6	0.6	241	2	US-08-353-476-116	Sequence 116, App
691	6	0.6	242	3	US-09-067-800-2	Sequence 2, Appli
692	6	0.6	242	3	US-09-105-652-2	Sequence 2, Appli
693	6	0.6	242	3	US-09-349-677-2	Sequence 2, Appli
694	6	0.6	242	4	US-09-708-584-2	Sequence 2, Appli
695	6	0.6	242	4	US-09-252-991A-19136	Sequence 19136, A
696	6	0.6	246	4	US-09-328-352-5613	Sequence 5613, Ap
697	6	0.6	248	2	US-08-851-974-3	Sequence 3, Appli
698	6	0.6	248	2	US-09-213-390-3	Sequence 3, Appli
699	6	0.6	248	3	US-08-750-145A-22	Sequence 22, Appl
700	6	0.6	248	3	US-08-975-698A-26	Sequence 26, Appl
701	6	0.6	248	3	US-09-417-090-26	Sequence 26, Appl
702	6	0.6	248	4	US-09-727-578-26	Sequence 26, Appl
703	6	0.6	248	4	US-09-252-991A-24788	Sequence 24788, A
704	6	0.6	251	1	US-08-209-747-8	Sequence 8, Appli
705	6	0.6	251	1	US-08-458-298-8	Sequence 8, Appli
706	6	0.6	251	3	US-08-875-811-59	Sequence 59, Appl
707	6	0.6	252	2	US-08-685-992-33	Sequence 33, Appl
708	6	0.6	252	2	US-09-144-925-33	Sequence 33, Appl
709	6	0.6	252	3	US-08-812-586-48	Sequence 48, Appl
710	6	0.6	252	4	US-09-489-847-214	Sequence 214, App
711	6	0.6	252	4	US-09-535-832A-45	Sequence 45, Appl
712	6	0.6	252	4	US-09-252-991A-26557	Sequence 26557, A
713	6	0.6	253	3	US-09-392-014-4	Sequence 4, Appli
714	6	0.6	253	4	US-09-328-352-6440	Sequence 6440, Ap
715	6	0.6	254	3	US-08-875-811-61	Sequence 61, Appl
716	6	0.6	254	4	US-09-134-001C-4565	Sequence 4565, Ap
717	6	0.6	254	4	US-08-914-479A-4	Sequence 4, Appli
718	6	0.6	254	4	US-09-252-991A-29819	Sequence 29819, A
719	6	0.6	255	4	US-08-858-207A-465	Sequence 465, App
720	6	0.6	255	4	US-09-107-532A-5335	Sequence 5335, Ap
721	6	0.6	256	2	US-09-008-960-3	Sequence 3, Appli
722	6	0.6	256	3	US-09-368-240-3	Sequence 3, Appli
723	6	0.6	256	4	US-09-468-702-3	Sequence 3, Appli
724	6	0.6	257	1	US-07-956-700B-105	Sequence 105, App
725	6	0.6	257	1	US-08-476-537-105	Sequence 105, App
726	6	0.6	257	1	US-08-485-607-105	Sequence 105, App
727	6	0.6	257	2	US-08-637-759B-92	Sequence 92, Appl
728	6	0.6	257	2	US-08-475-879-105	Sequence 105, App
729	6	0.6	257	3	US-08-871-355A-92	Sequence 92, Appl
730	6	0.6	257	4	US-09-201-945-92	Sequence 92, Appl
731	6	0.6	257	4	US-09-134-001C-3562	Sequence 3562, Ap
732	6	0.6	257	4	US-09-433-043B-105	Sequence 105, App
733	6	0.6	257	4	US-09-252-991A-26380	Sequence 26380, A
734	6	0.6	257	4	US-09-328-352-6621	Sequence 6621, Ap
735	6	0.6	257	4	US-09-107-532A-5440	Sequence 5440, Ap
736	6	0.6	258	1	US-07-721-761A-29	Sequence 29, Appl
737	6	0.6	258	1	US-07-978-687-29	Sequence 29, Appl
738	6	0.6	258	5	PCT-US91-05801-29	Sequence 29, Appl
739	6	0.6	260	4	US-09-252-991A-24755	Sequence 24755, A
740	6	0.6	261	4	US-09-489-847-371	Sequence 371, App
741	6	0.6	263	4	US-09-252-991A-27114	Sequence 27114, A
742	6	0.6	263	4	US-09-252-991A-27644	Sequence 27644, A
743	6	0.6	265	4	US-09-996-243-223	Sequence 223, App
744	6	0.6	268	3	US-09-353-585-6	Sequence 6, Appli
745	6	0.6	268	4	US-09-252-991A-17449	Sequence 17449, A
746	6	0.6	269	4	US-09-328-352-6139	Sequence 6139, Ap
747	6	0.6	270	4	US-09-205-258-568	Sequence 568, App
748	6	0.6	270	4	US-09-252-991A-30244	Sequence 30244, A
749	6	0.6	271	4	US-09-252-991A-17292	Sequence 17292, A
750	6	0.6	271	4	US-09-252-991A-23450	Sequence 23450, A
751	6	0.6	271	4	US-09-252-991A-32772	Sequence 32772, A
752	6	0.6	271	4	US-09-107-532A-5157	Sequence 5157, Ap
753	6	0.6	272	1	US-08-117-083-12	Sequence 12, Appl
754	6	0.6	272	2	US-08-709-874A-2	Sequence 2, Appli
755	6	0.6	272	3	US-09-104-382-2	Sequence 2, Appli
756	6	0.6	272	4	US-09-372-422A-26	Sequence 26, Appl
757	6	0.6	272	4	US-09-328-352-7372	Sequence 7372, Ap
758	6	0.6	273	2	US-08-353-476-112	Sequence 112, App
759	6	0.6	273	3	US-08-928-213B-14	Sequence 14, Appl
760	6	0.6	273	4	US-09-252-991A-22948	Sequence 22948, A
761	6	0.6	275	4	US-09-647-224A-26	Sequence 26, Appl
762	6	0.6	277	1	US-08-024-868-2	Sequence 2, Appli
763	6	0.6	277	2	US-08-890-542A-2	Sequence 2, Appli
764	6	0.6	277	2	US-08-242-097-2	Sequence 2, Appli
765	6	0.6	277	2	US-08-567-375-16	Sequence 16, Appl
766	6	0.6	277	3	US-08-591-605-2	Sequence 2, Appli
767	6	0.6	277	3	US-08-964-308-6	Sequence 6, Appli
768	6	0.6	277	3	US-08-964-308-10	Sequence 10, Appl
769	6	0.6	277	3	US-08-462-969B-4	Sequence 4, Appli
770	6	0.6	277	3	US-08-964-313-6	Sequence 6, Appli
771	6	0.6	277	3	US-08-964-313-10	Sequence 10, Appl
772	6	0.6	277	3	US-09-206-695-2	Sequence 2, Appli
773	6	0.6	277	4	US-09-000-179-1	Sequence 1, Appli
774	6	0.6	277	4	US-09-069-138-6	Sequence 6, Appli
775	6	0.6	277	4	US-09-069-138-10	Sequence 10, Appl
776	6	0.6	277	4	US-09-561-756-12	Sequence 12, Appl
777	6	0.6	277	4	US-09-227-721-12	Sequence 12, Appl
778	6	0.6	277	4	US-08-983-502-30	Sequence 30, Appl
779	6	0.6	277	4	US-09-124-934A-4	Sequence 4, Appli
780	6	0.6	277	4	US-08-724-378D-5	Sequence 5, Appli
781	6	0.6	277	4	US-09-799-118-2	Sequence 2, Appli
782	6	0.6	277	4	US-08-334-251D-4	Sequence 4, Appli
783	6	0.6	277	4	US-09-198-452A-10	Sequence 10, Appl
784	6	0.6	277	4	US-09-516-747-30	Sequence 30, Appl
785	6	0.6	277	5	PCT-US96-10521-30	Sequence 30, Appl
786	6	0.6	277	5	PCT-US96-11995-1	Sequence 1, Appli
787	6	0.6	278	1	US-08-188-582-9	Sequence 9, Appli
788	6	0.6	278	1	US-08-646-715-9	Sequence 9, Appli
789	6	0.6	279	1	US-08-312-387B-8	Sequence 8, Appli
790	6	0.6	279	1	US-08-683-426-8	Sequence 8, Appli
791	6	0.6	279	1	US-08-683-458-8	Sequence 8, Appli
792	6	0.6	279	2	US-08-878-360-8	Sequence 8, Appli
793	6	0.6	279	2	US-09-008-960-1	Sequence 1, Appli
794	6	0.6	279	3	US-09-368-240-1	Sequence 1, Appli
795	6	0.6	279	4	US-09-468-702-1	Sequence 1, Appli
796	6	0.6	279	4	US-09-333-412-8	Sequence 8, Appli
797	6	0.6	279	4	US-09-328-352-4229	Sequence 4229, Ap
798	6	0.6	279	5	PCT-US91-00899-8	Sequence 8, Appli
799	6	0.6	280	2	US-08-872-961A-5	Sequence 5, Appli
800	6	0.6	280	3	US-09-231-258-5	Sequence 5, Appli
801	6	0.6	280	4	US-09-194-146-10	Sequence 10, Appl
802	6	0.6	280	4	US-09-328-352-4975	Sequence 4975, Ap
803	6	0.6	282	2	US-08-284-941-8	Sequence 8, Appli
804	6	0.6	282	2	US-08-447-642-8	Sequence 8, Appli
805	6	0.6	282	3	US-09-236-503-8	Sequence 8, Appli
806	6	0.6	282	4	US-09-198-452A-739	Sequence 739, App
807	6	0.6	282	4	US-09-107-532A-5553	Sequence 5553, Ap
808	6	0.6	282	5	PCT-US93-02147A-8	Sequence 8, Appli
809	6	0.6	283	3	US-08-867-381A-3	Sequence 3, Appli
810	6	0.6	283	4	US-09-521-144-3	Sequence 3, Appli
811	6	0.6	283	4	US-09-134-001C-5346	Sequence 5346, Ap
812	6	0.6	283	4	US-09-134-001C-5491	Sequence 5491, Ap
813	6	0.6	283	4	US-09-609-816-8	Sequence 8, Appli
814	6	0.6	283	4	US-09-609-816-9	Sequence 9, Appli
815	6	0.6	284	4	US-08-914-479A-6	Sequence 6, Appli
816	6	0.6	285	4	US-09-561-756-35	Sequence 35, Appl
817	6	0.6	285	4	US-09-227-721-35	Sequence 35, Appl
818	6	0.6	285	4	US-09-107-532A-6136	Sequence 6136, Ap
819	6	0.6	287	2	US-08-424-641B-10	Sequence 10, Appl
820	6	0.6	287	2	US-08-820-980-10	Sequence 10, Appl
821	6	0.6	287	2	US-08-826-439-10	Sequence 10, Appl
822	6	0.6	287	4	US-09-252-991A-21605	Sequence 21605, A
823	6	0.6	287	4	US-09-252-991A-23091	Sequence 23091, A
824	6	0.6	287	4	US-09-252-991A-32232	Sequence 32232, A
825	6	0.6	288	2	US-08-147-772-2	Sequence 2, Appli
826	6	0.6	288	2	US-08-456-104-6	Sequence 6, Appli
827	6	0.6	288	2	US-08-101-624-23	Sequence 23, Appl
828	6	0.6	288	2	US-08-751-767A-6	Sequence 6, Appli
829	6	0.6	288	3	US-08-153-262-2	Sequence 2, Appli
830	6	0.6	288	3	US-08-479-744A-29	Sequence 29, Appl

831	6	0.6	288	3	US-08-280-757B-29	Sequence 29, Appl
832	6	0.6	288	3	US-09-159-135-2	Sequence 2, Appli
833	6	0.6	288	3	US-08-205-697A-19	Sequence 19, Appl
834	6	0.6	288	3	US-08-702-525-19	Sequence 19, Appl
835	6	0.6	288	4	US-09-450-798-2	Sequence 2, Appli
836	6	0.6	288	4	US-08-403-253A-2	Sequence 2, Appli
837	6	0.6	288	4	US-09-651-200-13	Sequence 13, Appl
838	6	0.6	288	4	US-09-651-200-14	Sequence 14, Appl
839	6	0.6	288	4	US-09-667-135-34	Sequence 34, Appl
840	6	0.6	288	4	US-08-435-816A-2	Sequence 2, Appli
841	6	0.6	288	4	US-09-107-532A-6783	Sequence 6783, Ap
842	6	0.6	288	5	PCT-US95-02576-19	Sequence 19, Appl
843	6	0.6	289	4	US-09-252-991A-20221	Sequence 20221, A
844	6	0.6	290	4	US-09-561-756-34	Sequence 34, Appl
845	6	0.6	290	4	US-09-227-721-34	Sequence 34, Appl
846	6	0.6	290	4	US-09-252-991A-18740	Sequence 18740, A
847	6	0.6	291	4	US-09-328-352-5269	Sequence 5269, Ap
848	6	0.6	292	2	US-09-024-848-2	Sequence 2, Appli
849	6	0.6	292	3	US-09-348-116A-2	Sequence 2, Appli
850	6	0.6	292	4	US-09-634-238-236	Sequence 236, App
851	6	0.6	292	5	PCT-US96-03916-5	Sequence 5, Appli
852	6	0.6	292	5	PCT-US96-03916-65	Sequence 65, Appl
853	6	0.6	292	6	5258287-24	Patent No. 5258287
854	6	0.6	293	4	US-09-428-747-2	Sequence 2, Appli
855	6	0.6	293	4	US-09-252-991A-26106	Sequence 26106, A
856	6	0.6	293	4	US-09-107-532A-5642	Sequence 5642, Ap
857	6	0.6	294	4	US-09-372-422A-40	Sequence 40, Appl
858	6	0.6	294	4	US-09-107-532A-5056	Sequence 5056, Ap
859	6	0.6	295	4	US-09-372-422A-38	Sequence 38, Appl
860	6	0.6	295	4	US-09-252-991A-32293	Sequence 32293, A
861	6	0.6	296	4	US-09-252-991A-27195	Sequence 27195, A
862	6	0.6	297	3	US-09-173-581-1	Sequence 1, Appli
863	6	0.6	297	3	US-09-420-915-1	Sequence 1, Appli
864	6	0.6	297	4	US-09-134-001C-5605	Sequence 5605, Ap
865	6	0.6	297	4	US-09-107-532A-6993	Sequence 6993, Ap
866	6	0.6	298	4	US-09-186-276B-30	Sequence 30, Appl
867	6	0.6	298	4	US-08-842-445-30	Sequence 30, Appl
868	6	0.6	298	4	US-09-186-188B-30	Sequence 30, Appl
869	6	0.6	299	2	US-08-709-874A-19	Sequence 19, Appl
870	6	0.6	299	3	US-09-104-382-19	Sequence 19, Appl
871	6	0.6	299	4	US-09-107-532A-4050	Sequence 4050, Ap
872	6	0.6	300	3	US-08-705-875A-6	Sequence 6, Appli
873	6	0.6	300	4	US-09-934-551-2	Sequence 2, Appli
874	6	0.6	300	4	US-09-242-999-6	Sequence 6, Appli
875	6	0.6	300	4	US-09-393-634-58	Sequence 58, Appl
876	6	0.6	301	2	US-08-716-301-2	Sequence 2, Appli
877	6	0.6	301	4	US-09-522-714-14	Sequence 14, Appl
878	6	0.6	302	4	US-09-252-991A-21900	Sequence 21900, A
879	6	0.6	303	3	US-08-867-381A-5	Sequence 5, Appli
880	6	0.6	303	4	US-09-521-144-5	Sequence 5, Appli
881	6	0.6	303	4	US-09-252-991A-22050	Sequence 22050, A
882	6	0.6	303	4	US-08-311-731A-93	Sequence 93, Appl
883	6	0.6	303	4	US-09-107-532A-6113	Sequence 6113, Ap
884	6	0.6	304	4	US-09-252-991A-31006	Sequence 31006, A
885	6	0.6	305	1	US-08-266-451B-26	Sequence 26, Appl
886	6	0.6	305	2	US-08-748-725-26	Sequence 26, Appl
887	6	0.6	305	3	US-08-937-271-10	Sequence 10, Appl
888	6	0.6	305	4	US-09-674-529B-14	Sequence 14, Appl
889	6	0.6	306	1	US-08-424-682A-1	Sequence 1, Appli
890	6	0.6	306	3	US-08-584-008A-8	Sequence 8, Appli
891	6	0.6	306	4	US-09-252-991A-22019	Sequence 22019, A
892	6	0.6	307	2	US-08-807-861A-56	Sequence 56, Appl
893	6	0.6	307	3	US-09-210-681-56	Sequence 56, Appl
894	6	0.6	307	3	US-08-946-719A-56	Sequence 56, Appl
895	6	0.6	307	4	US-09-547-983-56	Sequence 56, Appl
896	6	0.6	307	4	US-09-634-238-245	Sequence 245, App
897	6	0.6	307	4	US-09-252-991A-28821	Sequence 28821, A
898	6	0.6	307	4	US-09-107-532A-7040	Sequence 7040, Ap
899	6	0.6	308	3	US-08-705-875A-10	Sequence 10, Appl
900	6	0.6	308	4	US-09-242-999-10	Sequence 10, Appl
901	6	0.6	309	3	US-08-969-644-20	Sequence 20, Appl
902	6	0.6	309	3	US-08-444-189-20	Sequence 20, Appl
903	6	0.6	309	3	US-08-468-544-20	Sequence 20, Appl
904	6	0.6	309	4	US-09-724-623-72	Sequence 72, Appl
905	6	0.6	311	1	US-08-872-784-1	Sequence 1, Appli
906	6	0.6	311	2	US-08-686-599A-19	Sequence 19, Appl
907	6	0.6	311	2	US-09-100-851-1	Sequence 1, Appli
908	6	0.6	311	3	US-09-265-294-1	Sequence 1, Appli
909	6	0.6	311	4	US-09-252-991A-28792	Sequence 28792, A
910	6	0.6	311	4	US-09-328-352-5381	Sequence 5381, Ap
911	6	0.6	312	4	US-09-252-991A-19157	Sequence 19157, A
912	6	0.6	313	4	US-09-134-001C-4371	Sequence 4371, Ap
913	6	0.6	313	4	US-09-252-991A-28257	Sequence 28257, A
914	6	0.6	313	4	US-09-107-532A-6489	Sequence 6489, Ap
915	6	0.6	314	4	US-09-252-991A-18599	Sequence 18599, A
916	6	0.6	317	3	US-08-469-318-145	Sequence 145, App
917	6	0.6	317	3	US-08-468-609A-145	Sequence 145, App
918	6	0.6	317	4	US-08-446-872A-145	Sequence 145, App
919	6	0.6	317	4	US-08-762-227A-145	Sequence 145, App
920	6	0.6	317	4	US-09-252-991A-18820	Sequence 18820, A
921	6	0.6	317	5	PCT-US95-01185-145	Sequence 145, App
922	6	0.6	318	4	US-09-393-634-47	Sequence 47, Appl
923	6	0.6	319	5	PCT-US94-04208-2	Sequence 2, Appli
924	6	0.6	320	4	US-09-252-991A-18243	Sequence 18243, A
925	6	0.6	321	4	US-09-171-517B-12	Sequence 12, Appl
926	6	0.6	321	4	US-09-328-352-6676	Sequence 6676, Ap
927	6	0.6	321	4	US-09-107-532A-4118	Sequence 4118, Ap
928	6	0.6	322	4	US-09-107-532A-6135	Sequence 6135, Ap
929	6	0.6	323	4	US-09-252-991A-29849	Sequence 29849, A
930	6	0.6	325	3	US-08-867-381A-50	Sequence 50, Appl
931	6	0.6	325	4	US-09-521-144-50	Sequence 50, Appl
932	6	0.6	325	4	US-09-252-991A-20421	Sequence 20421, A
933	6	0.6	325	4	US-09-328-352-6196	Sequence 6196, Ap
934	6	0.6	326	4	US-09-134-001C-3274	Sequence 3274, Ap
935	6	0.6	326	4	US-09-252-991A-17002	Sequence 17002, A
936	6	0.6	326	4	US-09-252-991A-29943	Sequence 29943, A
937	6	0.6	326	4	US-09-198-452A-150	Sequence 150, App
938	6	0.6	327	3	US-09-392-014-2	Sequence 2, Appli
939	6	0.6	327	4	US-09-123-030-12	Sequence 12, Appl
940	6	0.6	328	1	US-08-265-087-2	Sequence 2, Appli
941	6	0.6	328	1	US-08-186-529-2	Sequence 2, Appli
942	6	0.6	328	1	US-08-621-493-2	Sequence 2, Appli
943	6	0.6	328	1	US-08-640-386A-2	Sequence 2, Appli
944	6	0.6	328	2	US-08-965-688-2	Sequence 2, Appli
945	6	0.6	328	2	US-08-751-767A-4	Sequence 4, Appli
946	6	0.6	328	3	US-09-260-173-2	Sequence 2, Appli
947	6	0.6	328	4	US-09-252-991A-25634	Sequence 25634, A
948	6	0.6	328	4	US-09-252-991A-30665	Sequence 30665, A
949	6	0.6	328	4	US-09-924-703-4	Sequence 4, Appli
950	6	0.6	330	2	US-08-815-176-1	Sequence 1, Appli
951	6	0.6	330	3	US-09-387-574-12	Sequence 12, Appl
952	6	0.6	330	4	US-09-668-096-12	Sequence 12, Appl
953	6	0.6	330	4	US-09-197-344-1	Sequence 1, Appli
954	6	0.6	330	4	US-09-535-909-5	Sequence 5, Appli
955	6	0.6	331	4	US-09-252-991A-19236	Sequence 19236, A
956	6	0.6	331	4	US-09-328-352-5339	Sequence 5339, Ap
957	6	0.6	332	4	US-09-252-991A-24826	Sequence 24826, A
958	6	0.6	333	4	US-09-252-991A-22893	Sequence 22893, A
959	6	0.6	333	4	US-09-328-352-7700	Sequence 7700, Ap
960	6	0.6	333	4	US-09-107-532A-4886	Sequence 4886, Ap
961	6	0.6	334	3	US-09-120-365-64	Sequence 64, Appl
962	6	0.6	334	3	US-09-515-039-64	Sequence 64, Appl
963	6	0.6	334	4	US-09-535-909-6	Sequence 6, Appli
964	6	0.6	335	2	US-08-405-175A-6	Sequence 6, Appli
965	6	0.6	335	4	US-09-252-991A-31163	Sequence 31163, A
966	6	0.6	336	4	US-09-252-991A-26226	Sequence 26226, A
967	6	0.6	336	4	US-09-252-991A-29645	Sequence 29645, A
968	6	0.6	337	4	US-09-328-352-5384	Sequence 5384, Ap
969	6	0.6	339	4	US-09-134-001C-3608	Sequence 3608, Ap
970	6	0.6	339	4	US-09-252-991A-28096	Sequence 28096, A
971	6	0.6	339	4	US-09-252-991A-30901	Sequence 30901, A
972	6	0.6	340	3	US-08-705-771-20	Sequence 20, Appl
973	6	0.6	340	4	US-09-651-200-2	Sequence 2, Appli
974	6	0.6	340	4	US-09-535-909-7	Sequence 7, Appli
975	6	0.6	341	4	US-08-311-731A-35	Sequence 35, Appl
976	6	0.6	344	4	US-09-198-452A-789	Sequence 789, App

977	6	0.6	344	4	US-09-198-452A-857	Sequence 857, App
978	6	0.6	344	6	5210183-2	Patent No. 5210183
979	6	0.6	345	2	US-08-479-614-5	Sequence 5, Appli
980	6	0.6	345	4	US-09-252-991A-21150	Sequence 21150, A
981	6	0.6	346	3	US-08-960-507-19	Sequence 19, Appl
982	6	0.6	346	4	US-09-136-801-19	Sequence 19, Appl
983	6	0.6	346	4	US-09-202-088A-19	Sequence 19, Appl
984	6	0.6	347	4	US-09-134-001C-5508	Sequence 5508, Ap
985	6	0.6	347	4	US-09-252-991A-18559	Sequence 18559, A
986	6	0.6	348	1	US-07-989-845-2	Sequence 2, Appli
987	6	0.6	348	5	PCT-US93-11298-2	Sequence 2, Appli
988	6	0.6	349	4	US-09-252-991A-24803	Sequence 24803, A
989	6	0.6	350	4	US-09-252-991A-18268	Sequence 18268, A
990	6	0.6	350	4	US-09-252-991A-29676	Sequence 29676, A
991	6	0.6	351	4	US-09-615-192A-397	Sequence 397, App
992	6	0.6	352	4	US-09-328-352-6481	Sequence 6481, Ap
993	6	0.6	353	4	US-09-134-001C-4325	Sequence 4325, Ap
994	6	0.6	353	4	US-09-292-858B-23	Sequence 23, Appl
995	6	0.6	353	4	US-09-252-991A-27528	Sequence 27528, A
996	6	0.6	354	2	US-08-700-013B-9	Sequence 9, Appli
997	6	0.6	354	4	US-09-134-001C-3899	Sequence 3899, Ap
998	6	0.6	354	4	US-09-198-452A-536	Sequence 536, App
999	6	0.6	354	4	US-09-328-352-4747	Sequence 4747, Ap
1000	6	0.6	354	4	US-09-328-352-7825	Sequence 7825, Ap

ALIGNMENTS

```
RESULT 1
US-08-258-188-2
; Sequence 2, Application US/08258188
; Patent No. 5475098
;  GENERAL INFORMATION:
;    APPLICANT:  HALL, Robert H.
;    APPLICANT:  XU, Jian Guo
;    TITLE OF INVENTION:  A NEW AND DISTINCTIVE DNA SEQUENCE OF E.
;    TITLE OF INVENTION:  coli O157:H7 AND ITS USE FOR THE RAPID, SENSITIVE AND
;    TITLE OF INVENTION:  SPECIFIC DETECTION OF O157:H7 AND OTHER ENTEROHEMORRHAGIC
;    TITLE OF INVENTION:  E. coli
;    NUMBER OF SEQUENCES:  29
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend Khourie and Crew
;      STREET:  Steuart Street Tower, One Market Plaza
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  US
;      ZIP:  94105-1493
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/258,188
;      FILING DATE:  14-JUN-1994
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Garrett-Wackowski, Eugenia
;      REGISTRATION NUMBER:  37,330
;      REFERENCE/DOCKET NUMBER:  15280206,DHHSE135940
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 543-9600
;      TELEFAX:  (415) 543-5043
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  758 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-258-188-2

  Query Match             1.5%;  Score 14;  DB 1;  Length 758;
  Best Local Similarity   100.0%;  Pred. No. 0.00022;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        379 ISGILEASKQAMFE 392
              ||||||||||||||
Db        152 ISGILEASKQAMFE 165


RESULT 2
US-08-526-813-2
; Sequence 2, Application US/08526813
; Patent No. 5756293
;  GENERAL INFORMATION:
;    APPLICANT:  Hall, Robert H.
;    APPLICANT:  Xu, Jian Guo
;    TITLE OF INVENTION:  A New and Distinctive DNA Sequence of E.
; Patent No. 5756293
;    TITLE OF INVENTION:  coli O157:H7 and its Use for the Rapid, Sensitive and
;    TITLE OF INVENTION:  Specific Detection of O157:H7 and Other Enterohemorrhagic
;    TITLE OF INVENTION:  E. coli
;    NUMBER OF SEQUENCES:  29
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend and Crew LLP
;      STREET:  Two Embarcadero Center, Eighth Floor
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94111-3834
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/526,813
;      FILING DATE:  11-SEP-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/258,188
;      FILING DATE:  14-JUN-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Garrett-Wackowski, Eugenia
;      REGISTRATION NUMBER:  37,330
;      REFERENCE/DOCKET NUMBER:  15280-206-1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 576-0200
;      TELEFAX:  (415) 576-0300
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  758 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-526-813-2

  Query Match             1.5%;  Score 14;  DB 1;  Length 758;
  Best Local Similarity   100.0%;  Pred. No. 0.00022;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        379 ISGILEASKQAMFE 392
              ||||||||||||||
Db        152 ISGILEASKQAMFE 165


RESULT 3
PCT-US95-08554-2
; Sequence 2, Application PC/TUS9508554
;  GENERAL INFORMATION:
;    APPLICANT:  HALL, Robert H.
;    APPLICANT:  XU, Jian Guo
;    TITLE OF INVENTION:  A NEW AND DISTINCTIVE DNA SEQUENCE OF E.
;    TITLE OF INVENTION:  coli O157:H7 AND ITS USE FOR THE RAPID, SENSITIVE AND
;    TITLE OF INVENTION:  SPECIFIC DETECTION OF O157:H7 AND OTHER ENTEROHEMORRHAGIC
```

```
;    TITLE OF INVENTION:  E. coli
;    NUMBER OF SEQUENCES:  29
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend Khourie and Crew
;      STREET:  Steuart Street Tower, One Market Plaza
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  US
;      ZIP:  94105-1493
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/08554
;      FILING DATE:  14-JUN-1994
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Garrett-Wackowski, Eugenia
;      REGISTRATION NUMBER:  37,330
;      REFERENCE/DOCKET NUMBER:  15280206,DHHSE135940
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 543-9600
;      TELEFAX:  (415) 543-5043
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  758 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
PCT-US95-08554-2

  Query Match             1.5%;  Score 14;  DB 5;  Length 758;
  Best Local Similarity   100.0%;  Pred. No. 0.00022;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        379 ISGILEASKQAMFE 392
              ||||||||||||||
Db        152 ISGILEASKQAMFE 165


RESULT 4
US-08-772-270A-11
; Sequence 11, Application US/08772270A
; Patent No. 6019984
;  GENERAL INFORMATION:
;    APPLICANT:  MacInnes, Janet
;    APPLICANT:  Ricciatti, Paul
;    APPLICANT:  Mallard, Bonnie
;    APPLICANT:  Rosendal, Soren
;    TITLE OF INVENTION:  NOVEL BACTERIAL PREPARATIONS, METHOD FOR
;    TITLE OF INVENTION:  PRODUCING SAME, AND THEIR USE AS VACCINES
;    NUMBER OF SEQUENCES:  14
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Bereskin & Parr
;      STREET:  40 King Street West
;      CITY:  Toronto
;      STATE:  Ontario
;      COUNTRY:  Canada
;      ZIP:  M5H 3Y2
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/772,270A
;      FILING DATE:  December 23, 1996
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Gravelle, Micheline
;      REGISTRATION NUMBER:  40,261
;      REFERENCE/DOCKET NUMBER:  6580-81
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (416) 364-7311
;      TELEFAX:  (416) 361-1398
;  INFORMATION FOR SEQ ID NO:  11:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1049 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    ORIGINAL SOURCE:
;      ORGANISM:  Actinobacillus pleuropneumoniae
US-08-772-270A-11

  Query Match             1.5%;  Score 14;  DB 3;  Length 1049;
  Best Local Similarity   100.0%;  Pred. No. 0.00029;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        241 KVAAGFELSNQVIG 254
              ||||||||||||||
Db        275 KVAAGFELSNQVIG 288


RESULT 5
US-09-062-126-10
; Sequence 10, Application US/09062126
; Patent No. 6500435
; GENERAL INFORMATION:
;  APPLICANT: Kamp, Elbarte Margriet
;  APPLICANT:  Smits, Marinus Adrianus
;  TITLE OF INVENTION: Recombinant Vaccine For  Prevention
;  TITLE OF INVENTION:  And/Or Treatment Of Pleuropneumonia Infections
;  FILE REFERENCE: 470-980537
;  CURRENT APPLICATION NUMBER: US/09/062,126
;  CURRENT FILING DATE:  1998-04-17
;  PRIOR APPLICATION NUMBER: 08/488,706
;  PRIOR FILING DATE: 1995-06-09
;  PRIOR APPLICATION NUMBER: 08/138,609
;  PRIOR FILING DATE: 1993-10-15
;  PRIOR APPLICATION NUMBER: 07/722,971
;  PRIOR FILING DATE: 1991-06-28
;  NUMBER OF SEQ ID NOS: 15
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 10
;   LENGTH: 1053
;   TYPE: PRT
;   ORGANISM: Actinobacillus pleuropneumonia
US-09-062-126-10

  Query Match             1.5%;  Score 14;  DB 4;  Length 1053;
  Best Local Similarity   100.0%;  Pred. No. 0.00029;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        241 KVAAGFELSNQVIG 254
              ||||||||||||||
Db        275 KVAAGFELSNQVIG 288


RESULT 6
PCT-US93-10500-2
; Sequence 2, Application PC/TUS9310500
;  GENERAL INFORMATION:
;    APPLICANT:  Chang, Yung-Fu
;    TITLE OF INVENTION:  Recombinant Vaccine For Procine
;    TITLE OF INVENTION:  Pleuropneumoniae
;    NUMBER OF SEQUENCES:  2
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Alan S. Korman
;      STREET:  1600 Empire Tower
;      CITY:  Buffalo
;      STATE:  New York
;      COUNTRY:  U.S.A.
```

```
;       ZIP:  14202
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  PCT/US93/10500
;       FILING DATE:
;       CLASSIFICATION:
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 07/972,229
;       FILING DATE:  05-NOV-1992
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Korman, Alan S.
;       REGISTRATION NUMBER:  33,932
;       REFERENCE/DOCKET NUMBER:  19603/00001
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  716-853-8104
;       TELEFAX:  716-853-8109
;   INFORMATION FOR SEQ ID NO:  2:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  1244 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  double
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  peptide
;     HYPOTHETICAL:  NO
;     ANTI-SENSE:  NO
;     FRAGMENT TYPE:  N-terminal
;     ORIGINAL SOURCE:
;       ORGANISM:  Actinobacillus pleuropneumoniae
;       STRAIN:  Serotypes 2, 3, 4, 6 and 8
;       INDIVIDUAL ISOLATE:  Swine
;       CELL TYPE:  Gram negative bacterium
PCT-US93-10500-2

  Query Match             1.5%;  Score 14;  DB 5;  Length 1244;
  Best Local Similarity   100.0%;  Pred. No. 0.00034;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        241 KVAAGFELSNQVIG 254
              ||||||||||||||
Db        448 KVAAGFELSNQVIG 461


RESULT 7
US-08-772-270A-8
; Sequence 8, Application US/08772270A
; Patent No. 6019984
;  GENERAL INFORMATION:
;     APPLICANT:  MacInnes, Janet
;     APPLICANT:  Ricciatti, Paul
;     APPLICANT:  Mallard, Bonnie
;     APPLICANT:  Rosendal, Soren
;     TITLE OF INVENTION:  NOVEL BACTERIAL PREPARATIONS, METHOD FOR
;     TITLE OF INVENTION:  PRODUCING SAME, AND THEIR USE AS VACCINES
;     NUMBER OF SEQUENCES:  14
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Bereskin & Parr
;       STREET:  40 King Street West
;       CITY:  Toronto
;       STATE:  Ontario
;       COUNTRY:  Canada
;       ZIP:  M5H 3Y2
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/772,270A
;       FILING DATE:  December 23, 1996
;       CLASSIFICATION:  424
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Gravelle, Micheline
;       REGISTRATION NUMBER:  40,261
;       REFERENCE/DOCKET NUMBER:  6580-81
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (416) 364-7311
;       TELEFAX:  (416) 361-1398
;   INFORMATION FOR SEQ ID NO:  8:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  956 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
;     ORIGINAL SOURCE:
;       ORGANISM:  Actinobacillus pleuropneumoniae
US-08-772-270A-8

  Query Match             1.2%;  Score 11;  DB 3;  Length 956;
  Best Local Similarity   100.0%;  Pred. No. 0.17;
  Matches   11;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        369 ALLVAGVTGLI 379
              |||||||||||
Db        392 ALLVAGVTGLI 402


RESULT 8
US-09-062-126-8
; Sequence 8, Application US/09062126
; Patent No. 6500435
; GENERAL INFORMATION:
;  APPLICANT: Kamp, Elbarte Margriet
;  APPLICANT:  Smits, Marinus Adrianus
;  TITLE OF INVENTION: Recombinant Vaccine For  Prevention
;  TITLE OF INVENTION:  And/Or Treatment Of Pleuropneumonia Infections
;  FILE REFERENCE: 470-980537
;  CURRENT APPLICATION NUMBER: US/09/062,126
;  CURRENT FILING DATE:  1998-04-17
;  PRIOR APPLICATION NUMBER: 08/488,706
;  PRIOR FILING DATE: 1995-06-09
;  PRIOR APPLICATION NUMBER: 08/138,609
;  PRIOR FILING DATE: 1993-10-15
;  PRIOR APPLICATION NUMBER: 07/722,971
;  PRIOR FILING DATE: 1991-06-28
;  NUMBER OF SEQ ID NOS: 15
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 8
;   LENGTH: 956
;   TYPE: PRT
;   ORGANISM: Actinobacillus pleuropneumonia
US-09-062-126-8

  Query Match             1.2%;  Score 11;  DB 4;  Length 956;
  Best Local Similarity   100.0%;  Pred. No. 0.17;
  Matches   11;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        369 ALLVAGVTGLI 379
              |||||||||||
Db        392 ALLVAGVTGLI 402


RESULT 9
US-09-062-126-3
; Sequence 3, Application US/09062126
; Patent No. 6500435
; GENERAL INFORMATION:
;  APPLICANT: Kamp, Elbarte Margriet
;  APPLICANT:  Smits, Marinus Adrianus
;  TITLE OF INVENTION: Recombinant Vaccine For  Prevention
;  TITLE OF INVENTION:  And/Or Treatment Of Pleuropneumonia Infections
;  FILE REFERENCE: 470-980537
```

```
;  CURRENT APPLICATION NUMBER: US/09/062,126
;  CURRENT FILING DATE:  1998-04-17
;  PRIOR APPLICATION NUMBER: 08/488,706
;  PRIOR FILING DATE: 1995-06-09
;  PRIOR APPLICATION NUMBER: 08/138,609
;  PRIOR FILING DATE: 1993-10-15
;  PRIOR APPLICATION NUMBER: 07/722,971
;  PRIOR FILING DATE: 1991-06-28
;  NUMBER OF SEQ ID NOS: 15
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 3
;   LENGTH: 1013
;   TYPE: PRT
;   ORGANISM: Actinobacillus pleuropneumonia
US-09-062-126-3

  Query Match             1.2%;  Score 11;  DB 4;  Length 1013;
  Best Local Similarity   100.0%;  Pred. No. 0.18;
  Matches   11;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        266 AQRVAAGLSTT 276
              |||||||||||
Db        289 AQRVAAGLSTT 299


RESULT 10
US-08-772-270A-2
; Sequence 2, Application US/08772270A
; Patent No. 6019984
;  GENERAL INFORMATION:
;    APPLICANT:  MacInnes, Janet
;    APPLICANT:  Ricciatti, Paul
;    APPLICANT:  Mallard, Bonnie
;    APPLICANT:  Rosendal, Soren
;    TITLE OF INVENTION:  NOVEL BACTERIAL PREPARATIONS, METHOD FOR
;    TITLE OF INVENTION:  PRODUCING SAME, AND THEIR USE AS VACCINES
;    NUMBER OF SEQUENCES:  14
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Bereskin & Parr
;      STREET:  40 King Street West
;      CITY:  Toronto
;      STATE:  Ontario
;      COUNTRY:  Canada
;      ZIP:  M5H 3Y2
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/772,270A
;      FILING DATE:  December 23, 1996
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Gravelle, Micheline
;      REGISTRATION NUMBER:  40,261
;      REFERENCE/DOCKET NUMBER:  6580-81
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (416) 364-7311
;      TELEFAX:  (416) 361-1398
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1022 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  circular
;    MOLECULE TYPE:  protein
;    ORIGINAL SOURCE:
;      ORGANISM:  Actinobacillus pleuropneumoniae
US-08-772-270A-2

  Query Match             1.2%;  Score 11;  DB 3;  Length 1022;
  Best Local Similarity   100.0%;  Pred. No. 0.18;
  Matches   11;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        266 AQRVAAGLSTT 276
              |||||||||||
Db        289 AQRVAAGLSTT 299


RESULT 11
US-08-387-942C-45
; Sequence 45, Application US/08387942C
; Patent No. 5939289
;  GENERAL INFORMATION:
;    APPLICANT:  ERTESVAG, HELGA
;    APPLICANT:  VALLA, SVEIN
;    APPLICANT:  SKJAK-BRAEK, GUDMUND
;    APPLICANT:  LARSEN, BJORN
;    TITLE OF INVENTION:  DNA COMPOUNDS COMPRISING SEQUENCES
;    TITLE OF INVENTION:  ENCODING MANNURONAN C-5-EPIMERASE
;    NUMBER OF SEQUENCES:  52
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  BIRCH, STEWART, KOLASCH & BIRCH, LLP
;      STREET:  P.O.BOX 747
;      CITY:  FALLS CHURCH
;      STATE:  VA
;      COUNTRY:  USA
;      ZIP:  22042
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/387,942C
;      FILING DATE:  09-MAY-1995
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  MURPHY JR, GERALD M.
;      REGISTRATION NUMBER:  28,977
;      REFERENCE/DOCKET NUMBER:  1809-106P
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  703-205-8000
;      TELEFAX:  703-205-8050
;  INFORMATION FOR SEQ ID NO:  45:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  149 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-387-942C-45

  Query Match             1.1%;  Score 10;  DB 2;  Length 149;
  Best Local Similarity   100.0%;  Pred. No. 0.27;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        714 GDDLLDGGAG 723
              ||||||||||
Db         31 GDDLLDGGAG 40


RESULT 12
US-09-252-149B-26
; Sequence 26, Application US/09252149B
; Patent No. 6369201
; GENERAL INFORMATION:
;  APPLICANT: Barker, Christopher A.
;  APPLICANT:  Morsey, Mohamad
;  TITLE OF INVENTION: IMMUNOLOGICAL METHODS TO MODULAR MYOSTATIN IN
;  TITLE OF INVENTION:  VERTEBRATE SUBJECTS
;  FILE REFERENCE: 9001-0042
;  CURRENT APPLICATION NUMBER: US/09/252,149B
;  CURRENT FILING DATE:  1999-02-18
;  PRIOR APPLICATION NUMBER: 60/075,213
```

```
;  PRIOR FILING DATE: 1998-02-19
;  NUMBER OF SEQ ID NOS: 39
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 26
;   LENGTH: 490
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: leukotoxin polypeptide carrie
;   OTHER INFORMATION: Figures 15A-15D
US-09-252-149B-26

  Query Match             1.1%;  Score 10;  DB 4;  Length 490;
  Best Local Similarity   100.0%;  Pred. No. 0.78;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        265 LAQRVAAGLS 274
              ||||||||||
Db        256 LAQRVAAGLS 265


RESULT 13
US-08-387-156-10
; Sequence 10, Application US/08387156
; Patent No. 5723129
;  GENERAL INFORMATION:
;    APPLICANT:  POTTER, ANDREW A.
;    APPLICANT:  REDMOND, MARK J.
;    APPLICANT:  HUGHES, HUW P.A.
;    TITLE OF INVENTION:  GnRH-LEUKOTOXIN CHIMERAS
;    NUMBER OF SEQUENCES:  28
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  REED & ROBINS
;      STREET:  635 BRYANT STREET
;      CITY:  PALO ALTO
;      STATE:  CALIFORNIA
;      COUNTRY:  UNITED STATES OF AMERICA
;      ZIP:  94301
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/387,156
;      FILING DATE:  10-FEB-1995
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/960,932
;      FILING DATE:  14-OCT-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/779,171
;      FILING DATE:  16-OCT-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  ROBINS, ROBERTA L.
;      REGISTRATION NUMBER:  33,208
;      REFERENCE/DOCKET NUMBER:  9001-0016.21
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 617-8999
;      TELEFAX:  (415) 327-3231
;  INFORMATION FOR SEQ ID NO:  10:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  544 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-387-156-10

  Query Match             1.1%;  Score 10;  DB 1;  Length 544;
  Best Local Similarity   100.0%;  Pred. No. 0.86;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        265 LAQRVAAGLS 274
              ||||||||||
Db        256 LAQRVAAGLS 265


RESULT 14
US-08-694-865-10
; Sequence 10, Application US/08694865
; Patent No. 5837268
;  GENERAL INFORMATION:
;    APPLICANT:  POTTER, ANDREW A.
;    APPLICANT:  MANNS, JOHN G.
;    TITLE OF INVENTION:  GnRH-LEUKOTOXIN CHIMERAS
;    NUMBER OF SEQUENCES:  34
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  REED & ROBINS LLP
;      STREET:  285 HAMILTON AVENUE, SUITE 200
;      CITY:  PALO ALTO
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94301
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/694,865
;      FILING DATE:  09-AUG-1996
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  MCCRACKEN, THOMAS P.
;      REGISTRATION NUMBER:  38,548
;      REFERENCE/DOCKET NUMBER:  9001-0016.22
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415)327-3400
;      TELEFAX:  (415)327-3231
;  INFORMATION FOR SEQ ID NO:  10:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  544 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-694-865-10

  Query Match             1.1%;  Score 10;  DB 2;  Length 544;
  Best Local Similarity   100.0%;  Pred. No. 0.86;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        265 LAQRVAAGLS 274
              ||||||||||
Db        256 LAQRVAAGLS 265


RESULT 15
US-08-878-748-10
; Sequence 10, Application US/08878748
; Patent No. 5969126
;  GENERAL INFORMATION:
;    APPLICANT:  POTTER, ANDREW A.
;    APPLICANT:  REDMOND, MARK J.
;    APPLICANT:  HUGHES, HUW P.A.
;    TITLE OF INVENTION:  GnRH-LEUKOTOXIN CHIMERAS
;    NUMBER OF SEQUENCES:  28
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  REED & ROBINS
;      STREET:  635 BRYANT STREET
;      CITY:  PALO ALTO
;      STATE:  CALIFORNIA
;      COUNTRY:  UNITED STATES OF AMERICA
;      ZIP:  94301
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
```

```
;     OPERATING SYSTEM:  PC-DOS/MS-DOS
;     SOFTWARE:  PatentIn Release #1.0, Version #1.25
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER:  US/08/878,748
;     FILING DATE:  19-JUN-1997
;     CLASSIFICATION:  536
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER:  US 08/387,156
;     FILING DATE:  10-FEB-1995
;     APPLICATION NUMBER:  US 07/960,932
;     FILING DATE:  14-OCT-1992
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER:  US 07/779,171
;     FILING DATE:  16-OCT-1991
;   ATTORNEY/AGENT INFORMATION:
;     NAME:  ROBINS, ROBERTA L.
;     REGISTRATION NUMBER:  33,208
;     REFERENCE/DOCKET NUMBER:  9001-0016.21
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE:  (415) 617-8999
;     TELEFAX:  (415) 327-3231
; INFORMATION FOR SEQ ID NO:  10:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH:  544 amino acids
;     TYPE:  amino acid
;     TOPOLOGY:  linear
;   MOLECULE TYPE:  protein
US-08-878-748-10

  Query Match             1.1%;  Score 10;  DB 2;  Length 544;
  Best Local Similarity   100.0%;  Pred. No. 0.86;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        265 LAQRVAAGLS 274
              ||||||||||
Db        256 LAQRVAAGLS 265


RESULT 16
US-09-124-491-10
; Sequence 10, Application US/09124491
; Patent No. 6022960
;  GENERAL INFORMATION:
;    APPLICANT:  POTTER, ANDREW A.
;    APPLICANT:  MANNS, JOHN G.
;    TITLE OF INVENTION:  GnRH-LEUKOTOXIN CHIMERAS
;    NUMBER OF SEQUENCES:  34
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  REED & ROBINS LLP
;      STREET:  285 HAMILTON AVENUE, SUITE 200
;      CITY:  PALO ALTO
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94301
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/124,491
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/694,865
;      FILING DATE:  09-AUG-1996
;      APPLICATION NUMBER:  US 08/387,156
;      FILING DATE:  10-FEB-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/960,932
;      FILING DATE:  14-OCT-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/779,171
;      FILING DATE:  16-OCT-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  MCCRACKEN, THOMAS P.
;      REGISTRATION NUMBER:  38,548
;      REFERENCE/DOCKET NUMBER:  9001-0016.22
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415)327-3400
;      TELEFAX:  (415)327-3231
;  INFORMATION FOR SEQ ID NO:  10:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  544 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-124-491-10

  Query Match             1.1%;  Score 10;  DB 3;  Length 544;
  Best Local Similarity   100.0%;  Pred. No. 0.86;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        265 LAQRVAAGLS 274
              ||||||||||
Db        256 LAQRVAAGLS 265


RESULT 17
US-09-383-912-10
; Sequence 10, Application US/09383912
; Patent No. 6521746
;  GENERAL INFORMATION:
;    APPLICANT:  POTTER, ANDREW A.
;    APPLICANT:  MANNS, JOHN G.
;    TITLE OF INVENTION:  GnRH-LEUKOTOXIN CHIMERAS
;    NUMBER OF SEQUENCES:  34
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  REED & ROBINS LLP
;      STREET:  285 HAMILTON AVENUE, SUITE 200
;      CITY:  PALO ALTO
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94301
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/383,912
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/694,865
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  MCCRACKEN, THOMAS P.
;      REGISTRATION NUMBER:  38,548
;      REFERENCE/DOCKET NUMBER:  9001-0016.22
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415)327-3400
;      TELEFAX:  (415)327-3231
;  INFORMATION FOR SEQ ID NO:  10:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  544 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-383-912-10

  Query Match             1.1%;  Score 10;  DB 4;  Length 544;
  Best Local Similarity   100.0%;  Pred. No. 0.86;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        265 LAQRVAAGLS 274
```

```
           ||||||||||
Db       256 LAQRVAAGLS 265

RESULT 18
US-08-694-865-16
; Sequence 16, Application US/08694865
; Patent No. 5837268
;  GENERAL INFORMATION:
;    APPLICANT:  POTTER, ANDREW A.
;    APPLICANT:  MANNS, JOHN G.
;    TITLE OF INVENTION:  GnRH-LEUKOTOXIN CHIMERAS
;    NUMBER OF SEQUENCES:  34
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  REED & ROBINS LLP
;      STREET:  285 HAMILTON AVENUE, SUITE 200
;      CITY:  PALO ALTO
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94301
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/694,865
;      FILING DATE:  09-AUG-1996
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  MCCRACKEN, THOMAS P.
;      REGISTRATION NUMBER:  38,548
;      REFERENCE/DOCKET NUMBER:  9001-0016.22
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415)327-3400
;      TELEFAX:  (415)327-3231
;  INFORMATION FOR SEQ ID NO:  16:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  699 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-694-865-16

  Query Match             1.1%;  Score 10;  DB 2;  Length 699;
  Best Local Similarity   100.0%;  Pred. No. 1.1;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       265 LAQRVAAGLS 274
             ||||||||||
Db       359 LAQRVAAGLS 368

RESULT 19
US-09-124-491-16
; Sequence 16, Application US/09124491
; Patent No. 6022960
;  GENERAL INFORMATION:
;    APPLICANT:  POTTER, ANDREW A.
;    APPLICANT:  MANNS, JOHN G.
;    TITLE OF INVENTION:  GnRH-LEUKOTOXIN CHIMERAS
;    NUMBER OF SEQUENCES:  34
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  REED & ROBINS LLP
;      STREET:  285 HAMILTON AVENUE, SUITE 200
;      CITY:  PALO ALTO
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94301
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/124,491
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/694,865
;      FILING DATE:  09-AUG-1996
;      APPLICATION NUMBER:  US 08/387,156
;      FILING DATE:  10-FEB-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/960,932
;      FILING DATE:  14-OCT-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/779,171
;      FILING DATE:  16-OCT-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  MCCRACKEN, THOMAS P.
;      REGISTRATION NUMBER:  38,548
;      REFERENCE/DOCKET NUMBER:  9001-0016.22
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415)327-3400
;      TELEFAX:  (415)327-3231
;  INFORMATION FOR SEQ ID NO:  16:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  699 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-124-491-16

  Query Match             1.1%;  Score 10;  DB 3;  Length 699;
  Best Local Similarity   100.0%;  Pred. No. 1.1;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       265 LAQRVAAGLS 274
             ||||||||||
Db       359 LAQRVAAGLS 368

RESULT 20
US-09-383-912-16
; Sequence 16, Application US/09383912
; Patent No. 6521746
;  GENERAL INFORMATION:
;    APPLICANT:  POTTER, ANDREW A.
;    APPLICANT:  MANNS, JOHN G.
;    TITLE OF INVENTION:  GnRH-LEUKOTOXIN CHIMERAS
;    NUMBER OF SEQUENCES:  34
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  REED & ROBINS LLP
;      STREET:  285 HAMILTON AVENUE, SUITE 200
;      CITY:  PALO ALTO
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94301
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/383,912
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/694,865
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  MCCRACKEN, THOMAS P.
;      REGISTRATION NUMBER:  38,548
;      REFERENCE/DOCKET NUMBER:  9001-0016.22
;    TELECOMMUNICATION INFORMATION:
```

```
;      TELEPHONE:  (415)327-3400
;      TELEFAX:  (415)327-3231
;  INFORMATION FOR SEQ ID NO:  16:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  699 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-383-912-16

  Query Match              1.1%;  Score 10;  DB 4;  Length 699;
  Best Local Similarity   100.0%;  Pred. No. 1.1;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        265 LAQRVAAGLS 274
              ||||||||||
Db        359 LAQRVAAGLS 368


RESULT 21
US-08-619-812-8
; Sequence 8, Application US/08619812
; Patent No. 6100066
;  GENERAL INFORMATION:
;    APPLICANT:  POTTER, ANDREW A.
;    APPLICANT:  THEISEN, MICHAEL
;    APPLICANT:  HARLAND, RICHARD J.
;    APPLICANT:  RIOUX, CLEMENT R.
;    TITLE OF INVENTION:  VACCINES FOR HAEMOPHILUS SOMNUS
;    NUMBER OF SEQUENCES:  10
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  REED & ROBINS
;      STREET:  635 BRYANT STREET
;      CITY:  PALO ALTO
;      STATE:  CALIFORNIA
;      COUNTRY:  UNITED STATES OF AMERICA
;      ZIP:  94301
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/619,812
;      FILING DATE:  15-MAR-1996
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/038,719
;      FILING DATE:  29-MAR-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  ROBINS, ROBERTA L.
;      REGISTRATION NUMBER:  33,208
;      REFERENCE/DOCKET NUMBER:  9000-0019.20
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 617-8999
;      TELEFAX:  (415) 327-3231
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  924 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-619-812-8

  Query Match              1.1%;  Score 10;  DB 3;  Length 924;
  Best Local Similarity   100.0%;  Pred. No. 1.4;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        265 LAQRVAAGLS 274
              ||||||||||
Db        256 LAQRVAAGLS 265
```

```
RESULT 22
US-07-908-253-2
; Sequence 2, Application US/07908253
; Patent No. 5534256
;  GENERAL INFORMATION:
;    APPLICANT:  POTTER, ANDREW A.
;    APPLICANT:  HARLAND, RICHARD J.
;    TITLE OF INVENTION:  HAEMOPHILUS SOMNUS OUTER MEMBRANE
;    TITLE OF INVENTION:  PROTEIN EXTRACT ENRICHED WITH IRON-REGULATED PROTEINS
;    NUMBER OF SEQUENCES:  6
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  ROBERTA L. ROBINS
;      STREET:  635 BRYANT STREET
;      CITY:  PALO ALTO
;      STATE:  CALIFORNIA
;      COUNTRY:  UNITED STATES OF AMERICA
;      ZIP:  94301
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/908,253
;      FILING DATE:  19920702
;      CLASSIFICATION:  420
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  ROBINS, ROBERTA L.
;      REGISTRATION NUMBER:  33,208
;      REFERENCE/DOCKET NUMBER:  9000-0026
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 617-8999
;      TELEFAX:  (415) 327-3231
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  926 amino acids
;      TYPE:  AMINO ACID
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-07-908-253-2

  Query Match              1.1%;  Score 10;  DB 1;  Length 926;
  Best Local Similarity   100.0%;  Pred. No. 1.4;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        265 LAQRVAAGLS 274
              ||||||||||
Db        256 LAQRVAAGLS 265


RESULT 23
US-08-455-970A-2
; Sequence 2, Application US/08455970A
; Patent No. 5708155
;  GENERAL INFORMATION:
;    APPLICANT:  POTTER, ANDREW A.
;    APPLICANT:  REDMOND, MARK J.
;    APPLICANT:  HUGHES, HUW P.A.
;    TITLE OF INVENTION:  ENHANCED IMMUNOGENICITY USING LEUKOTOXIN
;    TITLE OF INVENTION:  CHIMERAS
;    NUMBER OF SEQUENCES:  15
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  REED & ROBINS
;      STREET:  285 HAMILTON AVENUE, SUITE 200
;      CITY:  PALO ALTO
;      STATE:  CALIFORNIA
;      COUNTRY:  UNITED STATES OF AMERICA
;      ZIP:  94301
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
```

```
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/455,970A
;      FILING DATE:  31-MAY-1995
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/960,932
;      FILING DATE:  14-OCT-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  ROBINS, ROBERTA L.
;      REGISTRATION NUMBER:  33,208
;      REFERENCE/DOCKET NUMBER:  9001-0016.10
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 327-3400
;      TELEFAX:  (415) 327-3231
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  926 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-455-970A-2

  Query Match             1.1%;  Score 10;  DB 1;  Length 926;
  Best Local Similarity   100.0%;  Pred. No. 1.4;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        265 LAQRVAAGLS 274
              ||||||||||
Db        256 LAQRVAAGLS 265


RESULT 24
US-08-387-156-6
; Sequence 6, Application US/08387156
; Patent No. 5723129
;  GENERAL INFORMATION:
;    APPLICANT:  POTTER, ANDREW A.
;    APPLICANT:  REDMOND, MARK J.
;    APPLICANT:  HUGHES, HUW P.A.
;    TITLE OF INVENTION:  GnRH-LEUKOTOXIN CHIMERAS
;    NUMBER OF SEQUENCES:  28
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  REED & ROBINS
;      STREET:  635 BRYANT STREET
;      CITY:  PALO ALTO
;      STATE:  CALIFORNIA
;      COUNTRY:  UNITED STATES OF AMERICA
;      ZIP:  94301
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/387,156
;      FILING DATE:  10-FEB-1995
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/960,932
;      FILING DATE:  14-OCT-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/779,171
;      FILING DATE:  16-OCT-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  ROBINS, ROBERTA L.
;      REGISTRATION NUMBER:  33,208
;      REFERENCE/DOCKET NUMBER:  9001-0016.21
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 617-8999
;      TELEFAX:  (415) 327-3231
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  926 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-387-156-6

  Query Match             1.1%;  Score 10;  DB 1;  Length 926;
  Best Local Similarity   100.0%;  Pred. No. 1.4;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        265 LAQRVAAGLS 274
              ||||||||||
Db        256 LAQRVAAGLS 265


RESULT 25
US-08-694-865-6
; Sequence 6, Application US/08694865
; Patent No. 5837268
;  GENERAL INFORMATION:
;    APPLICANT:  POTTER, ANDREW A.
;    APPLICANT:  MANNS, JOHN G.
;    TITLE OF INVENTION:  GnRH-LEUKOTOXIN CHIMERAS
;    NUMBER OF SEQUENCES:  34
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  REED & ROBINS LLP
;      STREET:  285 HAMILTON AVENUE, SUITE 200
;      CITY:  PALO ALTO
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94301
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/694,865
;      FILING DATE:  09-AUG-1996
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  MCCRACKEN, THOMAS P.
;      REGISTRATION NUMBER:  38,548
;      REFERENCE/DOCKET NUMBER:  9001-0016.22
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415)327-3400
;      TELEFAX:  (415)327-3231
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  926 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-694-865-6

  Query Match             1.1%;  Score 10;  DB 2;  Length 926;
  Best Local Similarity   100.0%;  Pred. No. 1.4;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        265 LAQRVAAGLS 274
              ||||||||||
Db        256 LAQRVAAGLS 265


RESULT 26
US-08-878-748-6
; Sequence 6, Application US/08878748
; Patent No. 5969126
;  GENERAL INFORMATION:
;    APPLICANT:  POTTER, ANDREW A.
;    APPLICANT:  REDMOND, MARK J.
;    APPLICANT:  HUGHES, HUW P.A.
;    TITLE OF INVENTION:  GnRH-LEUKOTOXIN CHIMERAS
;    NUMBER OF SEQUENCES:  28
```

```
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  REED & ROBINS
;      STREET:  635 BRYANT STREET
;      CITY:  PALO ALTO
;      STATE:  CALIFORNIA
;      COUNTRY:  UNITED STATES OF AMERICA
;      ZIP:  94301
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/878,748
;      FILING DATE:  19-JUN-1997
;      CLASSIFICATION:  536
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/387,156
;      FILING DATE:  10-FEB-1995
;      APPLICATION NUMBER:  US 07/960,932
;      FILING DATE:  14-OCT-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/779,171
;      FILING DATE:  16-OCT-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  ROBINS, ROBERTA L.
;      REGISTRATION NUMBER:  33,208
;      REFERENCE/DOCKET NUMBER:  9001-0016.21
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 617-8999
;      TELEFAX:  (415) 327-3231
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  926 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-878-748-6

  Query Match             1.1%;  Score 10;  DB 2;  Length 926;
  Best Local Similarity   100.0%;  Pred. No. 1.4;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        265 LAQRVAAGLS 274
              ||||||||||
Db        256 LAQRVAAGLS 265


RESULT 27
US-08-535-837-2
; Sequence 2, Application US/08535837
; Patent No. 5985289
;  GENERAL INFORMATION:
;    APPLICANT:  POTTER, ANDREW A.
;    APPLICANT:  HARLAND, RICHARD J.
;    TITLE OF INVENTION:  HAEMOPHILUS SOMNUS OUTER MEMBRANE
;    TITLE OF INVENTION:  PROTEIN EXTRACT ENRICHED WITH IRON-REGULATED PROTEINS
;    NUMBER OF SEQUENCES:  6
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  ROBERTA L. ROBINS
;      STREET:  285 HAMILTON AVENUE, SUITE 200
;      CITY:  PALO ALTO
;      STATE:  CALIFORNIA
;      COUNTRY:  UNITED STATES OF AMERICA
;      ZIP:  94301
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/535,837
;      FILING DATE:  27-SEP-1995
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  ROBINS, ROBERTA L.
;      REGISTRATION NUMBER:  33,208
;      REFERENCE/DOCKET NUMBER:  9001-0026.01
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 327-3400
;      TELEFAX:  (415) 327-3231
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  926 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-535-837-2

  Query Match             1.1%;  Score 10;  DB 2;  Length 926;
  Best Local Similarity   100.0%;  Pred. No. 1.4;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        265 LAQRVAAGLS 274
              ||||||||||
Db        256 LAQRVAAGLS 265


RESULT 28
US-09-124-491-6
; Sequence 6, Application US/09124491
; Patent No. 6022960
;  GENERAL INFORMATION:
;    APPLICANT:  POTTER, ANDREW A.
;    APPLICANT:  MANNS, JOHN G.
;    TITLE OF INVENTION:  GnRH-LEUKOTOXIN CHIMERAS
;    NUMBER OF SEQUENCES:  34
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  REED & ROBINS LLP
;      STREET:  285 HAMILTON AVENUE, SUITE 200
;      CITY:  PALO ALTO
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94301
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/124,491
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/694,865
;      FILING DATE:  09-AUG-1996
;      APPLICATION NUMBER:  US 08/387,156
;      FILING DATE:  10-FEB-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/960,932
;      FILING DATE:  14-OCT-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/779,171
;      FILING DATE:  16-OCT-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  MCCRACKEN, THOMAS P.
;      REGISTRATION NUMBER:  38,548
;      REFERENCE/DOCKET NUMBER:  9001-0016.22
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415)327-3400
;      TELEFAX:  (415)327-3231
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  926 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
```

```
;    MOLECULE TYPE:  protein
US-09-124-491-6

  Query Match             1.1%;  Score 10;  DB 3;  Length 926;
  Best Local Similarity   100.0%;  Pred. No. 1.4;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        265 LAQRVAAGLS 274
              ||||||||||
Db        256 LAQRVAAGLS 265


RESULT 29
US-09-383-912-6
; Sequence 6, Application US/09383912
; Patent No. 6521746
;  GENERAL INFORMATION:
;    APPLICANT:  POTTER, ANDREW A.
;    APPLICANT:  MANNS, JOHN G.
;    TITLE OF INVENTION:  GnRH-LEUKOTOXIN CHIMERAS
;    NUMBER OF SEQUENCES:  34
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  REED & ROBINS LLP
;      STREET:  285 HAMILTON AVENUE, SUITE 200
;      CITY:  PALO ALTO
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94301
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/383,912
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/694,865
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  MCCRACKEN, THOMAS P.
;      REGISTRATION NUMBER:  38,548
;      REFERENCE/DOCKET NUMBER:  9001-0016.22
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415)327-3400
;      TELEFAX:  (415)327-3231
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  926 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-383-912-6

  Query Match             1.1%;  Score 10;  DB 4;  Length 926;
  Best Local Similarity   100.0%;  Pred. No. 1.4;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        265 LAQRVAAGLS 274
              ||||||||||
Db        256 LAQRVAAGLS 265


RESULT 30
5476657-3
;Patent No. 5476657
;    APPLICANT: POTTER, ANDREW A.
;    TITLE OF INVENTION: PASTEURELLA HAEMOLYTICA LEUKOTOXIN
;COMPOSITIONS AND USES THEREOF
;    NUMBER OF SEQUENCES: 8
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/15,537
;      FILING DATE: 09-FEB-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER: 504,850
;      FILING DATE: 05-APR-1990
;      APPLICATION NUMBER: 335,018
;      FILING DATE: 07-APR-1989
;SEQ ID NO:3:
;      LENGTH: 926
5476657-3

  Query Match             1.1%;  Score 10;  DB 6;  Length 926;
  Best Local Similarity   100.0%;  Pred. No. 1.4;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        265 LAQRVAAGLS 274
              ||||||||||
Db        256 LAQRVAAGLS 265


RESULT 31
US-08-215-805A-80
; Sequence 80, Application US/08215805A
; Patent No. 5559008
;  GENERAL INFORMATION:
;    APPLICANT:  Chang, Yung-Fu
;    TITLE OF INVENTION:  LEUKOTOXIN GENE FROM PASTEURELLA
;    TITLE OF INVENTION:  SUIS
;    NUMBER OF SEQUENCES:  84
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Nixon, Hargrave, Devans & Doyle
;      STREET:  Clinton Square, P.O. Box 1051
;      CITY:  Rochester
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  14603
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/215,805A
;      FILING DATE:  22-MAR-1994
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Timian, Susan J.
;      REGISTRATION NUMBER:  34,103
;      REFERENCE/DOCKET NUMBER:  19603/61 (D-1329A)
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (716) 263-1636
;      TELEFAX:  (716) 263-1600
;  INFORMATION FOR SEQ ID NO:  80:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  934 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    ORIGINAL SOURCE:
;      ORGANISM:  Pasteurella suis
;      STRAIN:  5943
;    IMMEDIATE SOURCE:
;      LIBRARY:  P. suis DNA in Bacteriophage lambda-dash
;      CLONE:  (Lambda)yfc33-37
US-08-215-805A-80

  Query Match             1.1%;  Score 10;  DB 1;  Length 934;
  Best Local Similarity   100.0%;  Pred. No. 1.4;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        265 LAQRVAAGLS 274
              ||||||||||
Db        279 LAQRVAAGLS 288
```

```
RESULT 32
US-08-455-970A-12
; Sequence 12, Application US/08455970A
; Patent No. 5708155
;  GENERAL INFORMATION:
;    APPLICANT:  POTTER, ANDREW A.
;    APPLICANT:  REDMOND, MARK J.
;    APPLICANT:  HUGHES, HUW P.A.
;    TITLE OF INVENTION:  ENHANCED IMMUNOGENICITY USING LEUKOTOXIN
;    TITLE OF INVENTION:  CHIMERAS
;    NUMBER OF SEQUENCES:  15
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  REED & ROBINS
;      STREET:  285 HAMILTON AVENUE, SUITE 200
;      CITY:  PALO ALTO
;      STATE:  CALIFORNIA
;      COUNTRY:  UNITED STATES OF AMERICA
;      ZIP:  94301
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/455,970A
;      FILING DATE:  31-MAY-1995
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/960,932
;      FILING DATE:  14-OCT-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  ROBINS, ROBERTA L.
;      REGISTRATION NUMBER:  33,208
;      REFERENCE/DOCKET NUMBER:  9001-0016.10
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 327-3400
;      TELEFAX:  (415) 327-3231
;  INFORMATION FOR SEQ ID NO:  12:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  936 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-455-970A-12

  Query Match             1.1%;  Score 10;  DB 1;  Length 936;
  Best Local Similarity   100.0%;  Pred. No. 1.4;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        265 LAQRVAAGLS 274
              ||||||||||
Db        256 LAQRVAAGLS 265


RESULT 33
US-08-455-970A-10
; Sequence 10, Application US/08455970A
; Patent No. 5708155
;  GENERAL INFORMATION:
;    APPLICANT:  POTTER, ANDREW A.
;    APPLICANT:  REDMOND, MARK J.
;    APPLICANT:  HUGHES, HUW P.A.
;    TITLE OF INVENTION:  ENHANCED IMMUNOGENICITY USING LEUKOTOXIN
;    TITLE OF INVENTION:  CHIMERAS
;    NUMBER OF SEQUENCES:  15
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  REED & ROBINS
;      STREET:  285 HAMILTON AVENUE, SUITE 200
;      CITY:  PALO ALTO
;      STATE:  CALIFORNIA
;      COUNTRY:  UNITED STATES OF AMERICA
;      ZIP:  94301
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/455,970A
;      FILING DATE:  31-MAY-1995
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/960,932
;      FILING DATE:  14-OCT-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  ROBINS, ROBERTA L.
;      REGISTRATION NUMBER:  33,208
;      REFERENCE/DOCKET NUMBER:  9001-0016.10
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 327-3400
;      TELEFAX:  (415) 327-3231
;  INFORMATION FOR SEQ ID NO:  10:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  943 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-455-970A-10

  Query Match             1.1%;  Score 10;  DB 1;  Length 943;
  Best Local Similarity   100.0%;  Pred. No. 1.4;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        265 LAQRVAAGLS 274
              ||||||||||
Db        256 LAQRVAAGLS 265


RESULT 34
US-08-455-970A-14
; Sequence 14, Application US/08455970A
; Patent No. 5708155
;  GENERAL INFORMATION:
;    APPLICANT:  POTTER, ANDREW A.
;    APPLICANT:  REDMOND, MARK J.
;    APPLICANT:  HUGHES, HUW P.A.
;    TITLE OF INVENTION:  ENHANCED IMMUNOGENICITY USING LEUKOTOXIN
;    TITLE OF INVENTION:  CHIMERAS
;    NUMBER OF SEQUENCES:  15
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  REED & ROBINS
;      STREET:  285 HAMILTON AVENUE, SUITE 200
;      CITY:  PALO ALTO
;      STATE:  CALIFORNIA
;      COUNTRY:  UNITED STATES OF AMERICA
;      ZIP:  94301
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/455,970A
;      FILING DATE:  31-MAY-1995
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/960,932
;      FILING DATE:  14-OCT-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  ROBINS, ROBERTA L.
;      REGISTRATION NUMBER:  33,208
;      REFERENCE/DOCKET NUMBER:  9001-0016.10
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 327-3400
```

```
;       TELEFAX:  (415) 327-3231
;   INFORMATION FOR SEQ ID NO:  14:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  951 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
US-08-455-970A-14

  Query Match              1.1%;  Score 10;  DB 1;  Length 951;
  Best Local Similarity   100.0%;  Pred. No. 1.4;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        265 LAQRVAAGLS 274
              ||||||||||
Db        256 LAQRVAAGLS 265


RESULT 35
US-08-387-156-8
; Sequence 8, Application US/08387156
; Patent No. 5723129
;  GENERAL INFORMATION:
;    APPLICANT:  POTTER, ANDREW A.
;    APPLICANT:  REDMOND, MARK J.
;    APPLICANT:  HUGHES, HUW P.A.
;    TITLE OF INVENTION:  GnRH-LEUKOTOXIN CHIMERAS
;    NUMBER OF SEQUENCES:  28
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  REED & ROBINS
;      STREET:  635 BRYANT STREET
;      CITY:  PALO ALTO
;      STATE:  CALIFORNIA
;      COUNTRY:  UNITED STATES OF AMERICA
;      ZIP:  94301
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/387,156
;      FILING DATE:  10-FEB-1995
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/960,932
;      FILING DATE:  14-OCT-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/779,171
;      FILING DATE:  16-OCT-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  ROBINS, ROBERTA L.
;      REGISTRATION NUMBER:  33,208
;      REFERENCE/DOCKET NUMBER:  9001-0016.21
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 617-8999
;      TELEFAX:  (415) 327-3231
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  977 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-387-156-8

  Query Match              1.1%;  Score 10;  DB 1;  Length 977;
  Best Local Similarity   100.0%;  Pred. No. 1.5;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        265 LAQRVAAGLS 274
              ||||||||||
Db        256 LAQRVAAGLS 265


RESULT 36
US-08-694-865-8
; Sequence 8, Application US/08694865
; Patent No. 5837268
;  GENERAL INFORMATION:
;    APPLICANT:  POTTER, ANDREW A.
;    APPLICANT:  MANNS, JOHN G.
;    TITLE OF INVENTION:  GnRH-LEUKOTOXIN CHIMERAS
;    NUMBER OF SEQUENCES:  34
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  REED & ROBINS LLP
;      STREET:  285 HAMILTON AVENUE, SUITE 200
;      CITY:  PALO ALTO
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94301
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/694,865
;      FILING DATE:  09-AUG-1996
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  MCCRACKEN, THOMAS P.
;      REGISTRATION NUMBER:  38,548
;      REFERENCE/DOCKET NUMBER:  9001-0016.22
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415)327-3400
;      TELEFAX:  (415)327-3231
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  977 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-694-865-8

  Query Match              1.1%;  Score 10;  DB 2;  Length 977;
  Best Local Similarity   100.0%;  Pred. No. 1.5;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        265 LAQRVAAGLS 274
              ||||||||||
Db        256 LAQRVAAGLS 265


RESULT 37
US-08-878-748-8
; Sequence 8, Application US/08878748
; Patent No. 5969126
;  GENERAL INFORMATION:
;    APPLICANT:  POTTER, ANDREW A.
;    APPLICANT:  REDMOND, MARK J.
;    APPLICANT:  HUGHES, HUW P.A.
;    TITLE OF INVENTION:  GnRH-LEUKOTOXIN CHIMERAS
;    NUMBER OF SEQUENCES:  28
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  REED & ROBINS
;      STREET:  635 BRYANT STREET
;      CITY:  PALO ALTO
;      STATE:  CALIFORNIA
;      COUNTRY:  UNITED STATES OF AMERICA
;      ZIP:  94301
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
```

```
;      APPLICATION NUMBER:  US/08/878,748
;      FILING DATE:  19-JUN-1997
;      CLASSIFICATION:  536
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/387,156
;      FILING DATE:  10-FEB-1995
;      APPLICATION NUMBER:  US 07/960,932
;      FILING DATE:  14-OCT-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/779,171
;      FILING DATE:  16-OCT-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  ROBINS, ROBERTA L.
;      REGISTRATION NUMBER:  33,208
;      REFERENCE/DOCKET NUMBER:  9001-0016.21
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 617-8999
;      TELEFAX:  (415) 327-3231
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  977 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-878-748-8

  Query Match              1.1%;  Score 10;  DB 2;  Length 977;
  Best Local Similarity   100.0%;  Pred. No. 1.5;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        265 LAQRVAAGLS 274
              ||||||||||
Db        256 LAQRVAAGLS 265


RESULT 38
US-09-124-491-8
; Sequence 8, Application US/09124491
; Patent No. 6022960
;  GENERAL INFORMATION:
;    APPLICANT:  POTTER, ANDREW A.
;    APPLICANT:  MANNS, JOHN G.
;    TITLE OF INVENTION:  GnRH-LEUKOTOXIN CHIMERAS
;    NUMBER OF SEQUENCES:  34
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  REED & ROBINS LLP
;      STREET:  285 HAMILTON AVENUE, SUITE 200
;      CITY:  PALO ALTO
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94301
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/124,491
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/694,865
;      FILING DATE:  09-AUG-1996
;      APPLICATION NUMBER:  US 08/387,156
;      FILING DATE:  10-FEB-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/960,932
;      FILING DATE:  14-OCT-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/779,171
;      FILING DATE:  16-OCT-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  MCCRACKEN, THOMAS P.
;      REGISTRATION NUMBER:  38,548
;      REFERENCE/DOCKET NUMBER:  9001-0016.22
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415)327-3400
;      TELEFAX:  (415)327-3231
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  977 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-124-491-8

  Query Match              1.1%;  Score 10;  DB 3;  Length 977;
  Best Local Similarity   100.0%;  Pred. No. 1.5;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        265 LAQRVAAGLS 274
              ||||||||||
Db        256 LAQRVAAGLS 265


RESULT 39
US-09-383-912-8
; Sequence 8, Application US/09383912
; Patent No. 6521746
;  GENERAL INFORMATION:
;    APPLICANT:  POTTER, ANDREW A.
;    APPLICANT:  MANNS, JOHN G.
;    TITLE OF INVENTION:  GnRH-LEUKOTOXIN CHIMERAS
;    NUMBER OF SEQUENCES:  34
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  REED & ROBINS LLP
;      STREET:  285 HAMILTON AVENUE, SUITE 200
;      CITY:  PALO ALTO
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94301
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/383,912
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/694,865
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  MCCRACKEN, THOMAS P.
;      REGISTRATION NUMBER:  38,548
;      REFERENCE/DOCKET NUMBER:  9001-0016.22
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415)327-3400
;      TELEFAX:  (415)327-3231
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  977 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-383-912-8

  Query Match              1.1%;  Score 10;  DB 4;  Length 977;
  Best Local Similarity   100.0%;  Pred. No. 1.5;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        265 LAQRVAAGLS 274
              ||||||||||
Db        256 LAQRVAAGLS 265
```

```
RESULT 40
US-07-777-715-9
; Sequence 9, Application US/07777715
; Patent No. 5273889
;  GENERAL INFORMATION:
;    APPLICANT:  Potter, Andrew
;    APPLICANT:  Campos, Manuel
;    APPLICANT:  Hughes, Huw P.A.
;    TITLE OF INVENTION:  CYTOKINE-LEUKOTOXIN GENE FUSIONS AND
;    TITLE OF INVENTION:  USES THEREOF
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Morrison & Foerster
;      STREET:  545 Middlefield Road, Suite 200
;      CITY:  Menlo Park
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94025
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/777,715
;      FILING DATE:  19911016
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Robins, Roberta L.
;      REGISTRATION NUMBER:  33,208
;      REFERENCE/DOCKET NUMBER:  29310-2001320
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-327-7250
;      TELEFAX:  415-327-2951
;      TELEX:  706141
;  INFORMATION FOR SEQ ID NO:  9:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1069 amino acids
;      TYPE:  AMINO ACID
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-07-777-715-9

  Query Match             1.1%;  Score 10;  DB 1;  Length 1069;
  Best Local Similarity   100.0%;  Pred. No. 1.6;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        265 LAQRVAAGLS 274
              ||||||||||
Db        256 LAQRVAAGLS 265


Search completed: February 17, 2004, 10:22:10
Job time : 33 secs
```

```
OM protein - protein search, using sw model

Run on:         February 17, 2004, 10:21:31 ; Search time 39 Seconds
                                              (without alignments)
                                              4976.856 Million cell updates/sec

Title:          US-10-069-799-5
Perfect score:  927
Sequence:       1 MSNINVIKSNIQAGLNSTKS..........SSNALQPITQPTQGILAPSV 927

Scoring table:  OLIGO
                Gapop 60.0 , Gapext 60.0

Searched:       801455 seqs, 209382283 residues

Word size :      0

Total number of hits satisfying chosen parameters:      801455

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database :       Published_Applications_AA:*
                 1:  /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
                 2:  /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
                 3:  /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
                 4:  /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
                 5:  /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
                 6:  /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
                 7:  /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
                 8:  /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
                 9:  /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
                 10:  /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
                 11:  /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
                 12:  /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
                 13:  /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
                 14:  /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
                 15:  /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
                 16:  /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
                 17:  /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
                 18:  /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.
```

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	826	89.1	927	11	US-09-884-696-2	Sequence 2, Appli
2	14	1.5	14	11	US-09-884-696-13	Sequence 13, Appl
3	14	1.5	1023	11	US-09-884-696-5	Sequence 5, Appli
4	12	1.3	12	11	US-09-884-696-6	Sequence 6, Appli
5	11	1.2	956	11	US-09-884-696-4	Sequence 4, Appli
6	10	1.1	695	11	US-09-305-924-13	Sequence 13, Appl
7	10	1.1	953	11	US-09-884-696-3	Sequence 3, Appli
8	10	1.1	1017	12	US-10-369-493-20166	Sequence 20166, A
9	9	1.0	259	15	US-10-156-761-12748	Sequence 12748, A
10	8	0.9	307	15	US-10-156-761-12788	Sequence 12788, A
11	8	0.9	408	15	US-10-156-761-14641	Sequence 14641, A
12	8	0.9	424	12	US-10-369-493-17978	Sequence 17978, A
13	8	0.9	477	12	US-10-369-493-18633	Sequence 18633, A
14	8	0.9	489	12	US-10-369-493-16555	Sequence 16555, A
15	8	0.9	505	12	US-10-369-493-17426	Sequence 17426, A
16	8	0.9	527	15	US-10-156-761-11340	Sequence 11340, A
17	8	0.9	773	12	US-10-369-493-1471	Sequence 1471, Ap
18	8	0.9	4327	12	US-10-369-493-10178	Sequence 10178, A
19	7	0.8	20	12	US-10-414-692-80	Sequence 80, Appl
20	7	0.8	44	8	US-08-424-550B-519	Sequence 519, App
21	7	0.8	49	12	US-10-264-049-3945	Sequence 3945, Ap
22	7	0.8	52	15	US-10-185-050-78	Sequence 78, Appl
23	7	0.8	63	9	US-09-917-340-9	Sequence 9, Appli
24	7	0.8	89	9	US-09-739-907-83	Sequence 83, Appl
25	7	0.8	89	12	US-09-938-671-83	Sequence 83, Appl
26	7	0.8	90	12	US-09-864-408A-5506	Sequence 5506, Ap
27	7	0.8	100	12	US-10-108-260A-3335	Sequence 3335, Ap
28	7	0.8	102	10	US-09-924-358-20	Sequence 20, Appl
29	7	0.8	102	12	US-10-410-764-20	Sequence 20, Appl
30	7	0.8	103	9	US-09-815-242-11108	Sequence 11108, A
31	7	0.8	104	9	US-09-741-669-297	Sequence 297, App
32	7	0.8	104	9	US-09-912-020-321	Sequence 321, App
33	7	0.8	104	9	US-09-815-242-10348	Sequence 10348, A
34	7	0.8	104	9	US-09-815-242-14107	Sequence 14107, A
35	7	0.8	104	12	US-10-287-274-323	Sequence 323, App
36	7	0.8	109	10	US-09-989-919-93	Sequence 93, Appl
37	7	0.8	114	9	US-09-864-761-40104	Sequence 40104, A
38	7	0.8	115	12	US-10-264-049-3972	Sequence 3972, Ap
39	7	0.8	116	12	US-10-094-749-2223	Sequence 2223, Ap
40	7	0.8	118	9	US-09-739-907-182	Sequence 182, App
41	7	0.8	118	12	US-09-938-671-182	Sequence 182, App
42	7	0.8	121	9	US-09-745-763-110	Sequence 110, App
43	7	0.8	122	12	US-10-312-187-52	Sequence 52, Appl
44	7	0.8	123	12	US-10-386-050A-8	Sequence 8, Appli
45	7	0.8	123	12	US-10-386-050A-10	Sequence 10, Appl
46	7	0.8	123	12	US-10-386-050A-22	Sequence 22, Appl
47	7	0.8	128	12	US-10-291-172-198	Sequence 198, App
48	7	0.8	134	9	US-09-925-301-1022	Sequence 1022, Ap
49	7	0.8	134	10	US-09-796-692-1206	Sequence 1206, Ap
50	7	0.8	134	10	US-09-796-692-1985	Sequence 1985, Ap
51	7	0.8	134	10	US-09-796-692-2226	Sequence 2226, Ap
52	7	0.8	134	12	US-10-057-475B-1206	Sequence 1206, Ap
53	7	0.8	134	12	US-10-057-475B-1985	Sequence 1985, Ap
54	7	0.8	134	12	US-10-057-475B-2226	Sequence 2226, Ap
55	7	0.8	134	12	US-10-154-884B-1206	Sequence 1206, Ap
56	7	0.8	134	12	US-10-154-884B-1985	Sequence 1985, Ap
57	7	0.8	134	12	US-10-154-884B-2226	Sequence 2226, Ap
58	7	0.8	134	15	US-10-040-862-1206	Sequence 1206, Ap
59	7	0.8	134	15	US-10-040-862-1985	Sequence 1985, Ap
60	7	0.8	134	15	US-10-040-862-2226	Sequence 2226, Ap
61	7	0.8	141	14	US-10-108-605-165	Sequence 165, App
62	7	0.8	151	12	US-10-376-564-58	Sequence 58, Appl
63	7	0.8	153	12	US-10-376-564-57	Sequence 57, Appl
64	7	0.8	155	12	US-10-094-749-2821	Sequence 2821, Ap
65	7	0.8	161	12	US-10-369-493-3946	Sequence 3946, Ap
66	7	0.8	174	12	US-10-264-237-2789	Sequence 2789, Ap
67	7	0.8	187	10	US-09-801-368-196	Sequence 196, App
68	7	0.8	187	12	US-10-369-493-22119	Sequence 22119, A
69	7	0.8	189	9	US-09-925-301-906	Sequence 906, App
70	7	0.8	190	10	US-09-764-868-862	Sequence 862, App
71	7	0.8	190	12	US-09-764-875-915	Sequence 915, App
72	7	0.8	190	12	US-09-764-875-1191	Sequence 1191, Ap
73	7	0.8	203	12	US-10-264-237-1513	Sequence 1513, Ap
74	7	0.8	209	12	US-10-427-442-20	Sequence 20, Appl
75	7	0.8	219	9	US-09-815-242-11973	Sequence 11973, A
76	7	0.8	219	15	US-10-205-823-313	Sequence 313, App
77	7	0.8	221	12	US-10-369-493-9552	Sequence 9552, Ap
78	7	0.8	235	12	US-10-029-386-32032	Sequence 32032, A
79	7	0.8	241	9	US-09-815-242-11653	Sequence 11653, A
80	7	0.8	246	12	US-10-369-493-11394	Sequence 11394, A
81	7	0.8	250	12	US-10-369-493-14790	Sequence 14790, A
82	7	0.8	252	12	US-10-369-493-19889	Sequence 19889, A
83	7	0.8	254	15	US-10-156-761-14194	Sequence 14194, A
84	7	0.8	256	9	US-09-815-242-10664	Sequence 10664, A
85	7	0.8	257	11	US-09-934-900-6	Sequence 6, Appli
86	7	0.8	260	15	US-10-156-761-13194	Sequence 13194, A
87	7	0.8	269	12	US-10-264-237-2168	Sequence 2168, Ap
88	7	0.8	274	12	US-09-953-348-125	Sequence 125, App

89	7	0.8	274	15	US-10-267-255-125	Sequence 125, App
90	7	0.8	275	9	US-09-925-302-737	Sequence 737, App
91	7	0.8	283	12	US-09-930-512-14	Sequence 14, Appl
92	7	0.8	284	12	US-10-114-828-18	Sequence 18, Appl
93	7	0.8	296	12	US-10-096-494-1	Sequence 1, Appli
94	7	0.8	296	12	US-10-369-493-8679	Sequence 8679, Ap
95	7	0.8	303	15	US-10-156-761-8055	Sequence 8055, Ap
96	7	0.8	308	10	US-09-738-626-6317	Sequence 6317, Ap
97	7	0.8	313	12	US-10-306-762-59	Sequence 59, Appl
98	7	0.8	320	9	US-09-854-122-20	Sequence 20, Appl
99	7	0.8	322	10	US-09-905-176-24	Sequence 24, Appl
100	7	0.8	324	9	US-09-815-242-11879	Sequence 11879, A
101	7	0.8	329	11	US-09-950-041-16	Sequence 16, Appl
102	7	0.8	333	10	US-09-976-059-2	Sequence 2, Appli
103	7	0.8	333	12	US-10-438-537-2	Sequence 2, Appli
104	7	0.8	338	12	US-10-369-493-16948	Sequence 16948, A
105	7	0.8	339	9	US-09-838-955-8	Sequence 8, Appli
106	7	0.8	339	15	US-10-156-761-9210	Sequence 9210, Ap
107	7	0.8	339	15	US-10-156-761-12600	Sequence 12600, A
108	7	0.8	341	12	US-10-369-493-18432	Sequence 18432, A
109	7	0.8	342	16	US-10-233-926-18	Sequence 18, Appl
110	7	0.8	343	15	US-10-156-761-9520	Sequence 9520, Ap
111	7	0.8	349	16	US-10-233-926-2	Sequence 2, Appli
112	7	0.8	349	16	US-10-233-926-12	Sequence 12, Appl
113	7	0.8	350	12	US-10-369-493-20851	Sequence 20851, A
114	7	0.8	353	15	US-10-106-698-6180	Sequence 6180, Ap
115	7	0.8	358	16	US-10-080-170-440	Sequence 440, App
116	7	0.8	359	12	US-10-369-493-20613	Sequence 20613, A
117	7	0.8	361	12	US-10-369-493-9259	Sequence 9259, Ap
118	7	0.8	361	12	US-10-369-493-9446	Sequence 9446, Ap
119	7	0.8	373	12	US-10-264-049-3049	Sequence 3049, Ap
120	7	0.8	374	12	US-10-369-493-17585	Sequence 17585, A
121	7	0.8	379	12	US-10-158-034-65	Sequence 65, Appl
122	7	0.8	379	12	US-10-158-034-90	Sequence 90, Appl
123	7	0.8	380	11	US-09-934-900-12	Sequence 12, Appl
124	7	0.8	380	12	US-10-411-976-9	Sequence 9, Appli
125	7	0.8	382	10	US-09-881-752A-224	Sequence 224, App
126	7	0.8	382	14	US-10-078-929-34	Sequence 34, Appl
127	7	0.8	385	15	US-10-156-761-12542	Sequence 12542, A
128	7	0.8	386	10	US-09-881-752A-74	Sequence 74, Appl
129	7	0.8	386	11	US-09-988-067B-32	Sequence 32, Appl
130	7	0.8	386	12	US-10-369-493-12670	Sequence 12670, A
131	7	0.8	386	15	US-10-204-887-142	Sequence 142, App
132	7	0.8	388	12	US-10-369-493-19947	Sequence 19947, A
133	7	0.8	389	12	US-10-369-493-19972	Sequence 19972, A
134	7	0.8	389	14	US-10-036-507-13	Sequence 13, Appl
135	7	0.8	390	12	US-10-369-493-18060	Sequence 18060, A
136	7	0.8	390	12	US-10-369-493-19694	Sequence 19694, A
137	7	0.8	393	12	US-10-369-493-20721	Sequence 20721, A
138	7	0.8	393	12	US-10-374-780A-1937	Sequence 1937, Ap
139	7	0.8	396	12	US-10-084-846A-76	Sequence 76, Appl
140	7	0.8	400	10	US-09-373-967-2	Sequence 2, Appli
141	7	0.8	400	12	US-10-369-493-17487	Sequence 17487, A
142	7	0.8	407	12	US-10-166-225A-84	Sequence 84, Appl
143	7	0.8	424	10	US-09-843-905A-13	Sequence 13, Appl
144	7	0.8	424	12	US-10-317-250-13	Sequence 13, Appl
145	7	0.8	435	12	US-09-813-432-48	Sequence 48, Appl
146	7	0.8	435	12	US-10-174-364-48	Sequence 48, Appl
147	7	0.8	436	12	US-10-369-493-21209	Sequence 21209, A
148	7	0.8	440	11	US-09-910-186A-8	Sequence 8, Appli
149	7	0.8	441	12	US-10-452-024-188	Sequence 188, App
150	7	0.8	442	12	US-10-369-493-20368	Sequence 20368, A
151	7	0.8	449	12	US-10-369-493-2938	Sequence 2938, Ap
152	7	0.8	449	12	US-10-369-493-21458	Sequence 21458, A
153	7	0.8	450	9	US-09-795-693-34	Sequence 34, Appl
154	7	0.8	450	12	US-10-302-267-96	Sequence 96, Appl
155	7	0.8	450	12	US-10-374-780A-2128	Sequence 2128, Ap
156	7	0.8	450	15	US-10-156-239-34	Sequence 34, Appl
157	7	0.8	450	15	US-10-199-485-34	Sequence 34, Appl
158	7	0.8	454	12	US-10-238-075-1220	Sequence 1220, Ap
159	7	0.8	454	12	US-10-369-493-1302	Sequence 1302, Ap
160	7	0.8	454	12	US-10-369-493-21598	Sequence 21598, A
161	7	0.8	455	9	US-09-815-242-11558	Sequence 11558, A
162	7	0.8	457	12	US-10-369-493-9323	Sequence 9323, Ap
163	7	0.8	458	9	US-09-815-242-11321	Sequence 11321, A
164	7	0.8	460	11	US-09-759-595-1	Sequence 1, Appli
165	7	0.8	460	15	US-10-001-254-16	Sequence 16, Appl
166	7	0.8	460	15	US-10-001-254-28	Sequence 28, Appl
167	7	0.8	463	15	US-10-156-761-14014	Sequence 14014, A
168	7	0.8	468	11	US-09-746-660A-66	Sequence 66, Appl
169	7	0.8	472	12	US-10-354-774-46	Sequence 46, Appl
170	7	0.8	472	12	US-10-271-012-46	Sequence 46, Appl
171	7	0.8	477	12	US-10-369-493-17633	Sequence 17633, A
172	7	0.8	478	15	US-10-156-761-11009	Sequence 11009, A
173	7	0.8	481	12	US-10-369-493-10573	Sequence 10573, A
174	7	0.8	493	12	US-10-369-493-23665	Sequence 23665, A
175	7	0.8	494	12	US-10-369-493-12924	Sequence 12924, A
176	7	0.8	499	12	US-10-369-493-20175	Sequence 20175, A
177	7	0.8	516	12	US-10-369-493-16571	Sequence 16571, A
178	7	0.8	519	15	US-10-205-823-299	Sequence 299, App
179	7	0.8	521	10	US-09-884-566-2	Sequence 2, Appli
180	7	0.8	522	15	US-10-156-761-12853	Sequence 12853, A
181	7	0.8	523	15	US-10-219-220-183	Sequence 183, App
182	7	0.8	551	10	US-09-738-626-4431	Sequence 4431, Ap
183	7	0.8	555	12	US-10-369-493-17959	Sequence 17959, A
184	7	0.8	585	9	US-09-815-242-11132	Sequence 11132, A
185	7	0.8	587	10	US-09-815-923-4	Sequence 4, Appli
186	7	0.8	587	12	US-10-238-075-1121	Sequence 1121, Ap
187	7	0.8	594	12	US-10-320-800-54	Sequence 54, Appl
188	7	0.8	595	12	US-10-369-493-5268	Sequence 5268, Ap
189	7	0.8	600	12	US-10-369-493-19505	Sequence 19505, A
190	7	0.8	605	12	US-10-104-047-3739	Sequence 3739, Ap
191	7	0.8	617	12	US-10-354-358-118	Sequence 118, App
192	7	0.8	622	9	US-09-843-598-11	Sequence 11, Appl
193	7	0.8	622	12	US-10-369-493-21226	Sequence 21226, A
194	7	0.8	630	12	US-10-369-493-421	Sequence 421, App
195	7	0.8	630	12	US-10-369-493-14775	Sequence 14775, A
196	7	0.8	634	12	US-10-369-493-11753	Sequence 11753, A
197	7	0.8	635	12	US-10-369-493-14237	Sequence 14237, A
198	7	0.8	635	12	US-10-369-493-14928	Sequence 14928, A
199	7	0.8	636	12	US-10-108-260A-4790	Sequence 4790, Ap
200	7	0.8	637	12	US-10-369-493-12290	Sequence 12290, A
201	7	0.8	640	12	US-10-369-493-20270	Sequence 20270, A
202	7	0.8	641	12	US-10-262-666-8	Sequence 8, Appli
203	7	0.8	655	12	US-10-369-493-12943	Sequence 12943, A
204	7	0.8	660	9	US-09-841-132-446	Sequence 446, App
205	7	0.8	662	12	US-10-369-493-5336	Sequence 5336, Ap
206	7	0.8	674	12	US-10-291-265-471	Sequence 471, App
207	7	0.8	698	12	US-10-369-493-20655	Sequence 20655, A
208	7	0.8	708	12	US-10-369-493-15900	Sequence 15900, A
209	7	0.8	708	12	US-10-369-493-16271	Sequence 16271, A
210	7	0.8	710	9	US-09-815-242-5367	Sequence 5367, Ap
211	7	0.8	711	9	US-09-815-242-12190	Sequence 12190, A
212	7	0.8	714	15	US-10-156-761-11928	Sequence 11928, A
213	7	0.8	732	11	US-09-977-418-20	Sequence 20, Appl
214	7	0.8	732	11	US-09-977-033A-20	Sequence 20, Appl
215	7	0.8	732	12	US-09-977-751C-20	Sequence 20, Appl
216	7	0.8	732	12	US-09-977-639A-20	Sequence 20, Appl
217	7	0.8	732	12	US-09-977-819B-20	Sequence 20, Appl
218	7	0.8	735	12	US-10-262-666-66	Sequence 66, Appl
219	7	0.8	744	12	US-10-262-666-64	Sequence 64, Appl
220	7	0.8	756	12	US-10-369-493-16891	Sequence 16891, A
221	7	0.8	781	12	US-10-187-253A-27	Sequence 27, Appl
222	7	0.8	781	12	US-10-187-253A-33	Sequence 33, Appl
223	7	0.8	808	15	US-10-156-761-10993	Sequence 10993, A
224	7	0.8	810	15	US-10-156-761-9208	Sequence 9208, Ap
225	7	0.8	812	15	US-10-210-296-6	Sequence 6, Appli
226	7	0.8	827	12	US-10-369-493-21772	Sequence 21772, A
227	7	0.8	828	12	US-10-425-913-3	Sequence 3, Appli
228	7	0.8	829	12	US-10-161-927-86	Sequence 86, Appl
229	7	0.8	847	12	US-10-369-493-6478	Sequence 6478, Ap
230	7	0.8	864	12	US-10-369-493-16631	Sequence 16631, A
231	7	0.8	893	9	US-09-916-790-5	Sequence 5, Appli
232	7	0.8	907	12	US-10-369-493-15529	Sequence 15529, A
233	7	0.8	1000	11	US-09-823-187-6	Sequence 6, Appli
234	7	0.8	1032	11	US-09-950-041-37	Sequence 37, Appl

235	7	0.8	1034	9	US-09-815-242-10331	Sequence 10331, A
236	7	0.8	1036	10	US-09-373-967-4	Sequence 4, Appli
237	7	0.8	1036	11	US-09-887-527-60	Sequence 60, Appl
238	7	0.8	1036	12	US-10-137-870-142	Sequence 142, App
239	7	0.8	1036	12	US-10-140-018-142	Sequence 142, App
240	7	0.8	1036	12	US-10-140-021-142	Sequence 142, App
241	7	0.8	1036	12	US-10-140-274-142	Sequence 142, App
242	7	0.8	1036	12	US-10-140-471-142	Sequence 142, App
243	7	0.8	1036	12	US-10-140-807-142	Sequence 142, App
244	7	0.8	1036	12	US-10-140-922-142	Sequence 142, App
245	7	0.8	1036	12	US-10-140-924-142	Sequence 142, App
246	7	0.8	1036	12	US-10-140-926-142	Sequence 142, App
247	7	0.8	1036	12	US-10-141-698-142	Sequence 142, App
248	7	0.8	1036	12	US-10-141-702-142	Sequence 142, App
249	7	0.8	1036	12	US-10-141-704-142	Sequence 142, App
250	7	0.8	1036	12	US-10-142-421-142	Sequence 142, App
251	7	0.8	1036	12	US-10-142-432-142	Sequence 142, App
252	7	0.8	1036	12	US-10-142-767-142	Sequence 142, App
253	7	0.8	1036	12	US-10-143-033-142	Sequence 142, App
254	7	0.8	1036	12	US-10-144-994-142	Sequence 142, App
255	7	0.8	1036	12	US-10-145-628-142	Sequence 142, App
256	7	0.8	1036	12	US-10-145-631-142	Sequence 142, App
257	7	0.8	1036	12	US-10-145-633-142	Sequence 142, App
258	7	0.8	1036	12	US-10-145-746-142	Sequence 142, App
259	7	0.8	1036	12	US-10-145-748-142	Sequence 142, App
260	7	0.8	1036	12	US-10-145-823-142	Sequence 142, App
261	7	0.8	1036	12	US-10-145-826-142	Sequence 142, App
262	7	0.8	1036	12	US-10-145-870-142	Sequence 142, App
263	7	0.8	1036	12	US-10-145-876-142	Sequence 142, App
264	7	0.8	1036	12	US-10-145-959-142	Sequence 142, App
265	7	0.8	1036	12	US-10-146-724-142	Sequence 142, App
266	7	0.8	1036	12	US-10-146-725-142	Sequence 142, App
267	7	0.8	1036	12	US-10-146-795-142	Sequence 142, App
268	7	0.8	1036	12	US-10-147-495-142	Sequence 142, App
269	7	0.8	1036	12	US-10-147-501-142	Sequence 142, App
270	7	0.8	1036	12	US-10-147-504-142	Sequence 142, App
271	7	0.8	1036	12	US-10-147-506-142	Sequence 142, App
272	7	0.8	1036	12	US-10-147-509-142	Sequence 142, App
273	7	0.8	1036	12	US-10-147-510-142	Sequence 142, App
274	7	0.8	1036	12	US-10-147-511-142	Sequence 142, App
275	7	0.8	1036	12	US-10-147-529-142	Sequence 142, App
276	7	0.8	1036	12	US-10-152-397-142	Sequence 142, App
277	7	0.8	1036	12	US-10-153-586-142	Sequence 142, App
278	7	0.8	1036	12	US-10-158-783-142	Sequence 142, App
279	7	0.8	1036	12	US-10-158-786-142	Sequence 142, App
280	7	0.8	1036	12	US-10-140-019-142	Sequence 142, App
281	7	0.8	1036	12	US-10-140-022-142	Sequence 142, App
282	7	0.8	1036	12	US-10-140-861-142	Sequence 142, App
283	7	0.8	1036	12	US-10-140-862-142	Sequence 142, App
284	7	0.8	1036	12	US-10-141-697-142	Sequence 142, App
285	7	0.8	1036	12	US-10-141-700-142	Sequence 142, App
286	7	0.8	1036	12	US-10-141-705-142	Sequence 142, App
287	7	0.8	1036	12	US-10-141-753-142	Sequence 142, App
288	7	0.8	1036	12	US-10-141-758-142	Sequence 142, App
289	7	0.8	1036	12	US-10-142-418-142	Sequence 142, App
290	7	0.8	1036	12	US-10-142-420-142	Sequence 142, App
291	7	0.8	1036	12	US-10-142-422-142	Sequence 142, App
292	7	0.8	1036	12	US-10-142-427-142	Sequence 142, App
293	7	0.8	1036	12	US-10-142-760-142	Sequence 142, App
294	7	0.8	1036	12	US-10-145-821-142	Sequence 142, App
295	7	0.8	1036	12	US-10-152-531-142	Sequence 142, App
296	7	0.8	1036	12	US-10-127-840A-142	Sequence 142, App
297	7	0.8	1036	12	US-10-142-424-142	Sequence 142, App
298	7	0.8	1036	12	US-10-142-761-142	Sequence 142, App
299	7	0.8	1036	12	US-10-142-763-142	Sequence 142, App
300	7	0.8	1036	12	US-10-142-765-142	Sequence 142, App
301	7	0.8	1036	12	US-10-142-887-142	Sequence 142, App
302	7	0.8	1036	12	US-10-142-888-142	Sequence 142, App
303	7	0.8	1036	12	US-10-143-034-142	Sequence 142, App
304	7	0.8	1036	12	US-10-143-116-142	Sequence 142, App
305	7	0.8	1036	12	US-10-144-957-142	Sequence 142, App
306	7	0.8	1036	12	US-10-144-992-142	Sequence 142, App
307	7	0.8	1036	12	US-10-145-015-142	Sequence 142, App

308	7	0.8	1036	12	US-10-145-090-142	Sequence 142, App
309	7	0.8	1036	12	US-10-145-091-142	Sequence 142, App
310	7	0.8	1036	12	US-10-145-629-142	Sequence 142, App
311	7	0.8	1036	12	US-10-145-630-142	Sequence 142, App
312	7	0.8	1036	12	US-10-145-747-142	Sequence 142, App
313	7	0.8	1036	12	US-10-145-752-142	Sequence 142, App
314	7	0.8	1036	12	US-10-145-754-142	Sequence 142, App
315	7	0.8	1036	12	US-10-145-755-142	Sequence 142, App
316	7	0.8	1036	12	US-10-145-818-142	Sequence 142, App
317	7	0.8	1036	12	US-10-145-820-142	Sequence 142, App
318	7	0.8	1036	12	US-10-145-872-142	Sequence 142, App
319	7	0.8	1036	12	US-10-145-873-142	Sequence 142, App
320	7	0.8	1036	12	US-10-147-481-142	Sequence 142, App
321	7	0.8	1036	12	US-10-147-482-142	Sequence 142, App
322	7	0.8	1036	12	US-10-147-503-142	Sequence 142, App
323	7	0.8	1036	12	US-10-147-522-142	Sequence 142, App
324	7	0.8	1036	12	US-10-152-401-142	Sequence 142, App
325	7	0.8	1036	12	US-10-157-783-142	Sequence 142, App
326	7	0.8	1036	12	US-10-158-462-142	Sequence 142, App
327	7	0.8	1036	12	US-10-158-792-142	Sequence 142, App
328	7	0.8	1036	12	US-10-143-035-142	Sequence 142, App
329	7	0.8	1036	12	US-10-145-751-142	Sequence 142, App
330	7	0.8	1036	12	US-10-145-822-142	Sequence 142, App
331	7	0.8	1036	12	US-10-145-824-142	Sequence 142, App
332	7	0.8	1036	12	US-10-145-827-142	Sequence 142, App
333	7	0.8	1036	12	US-10-145-869-142	Sequence 142, App
334	7	0.8	1036	12	US-10-145-875-142	Sequence 142, App
335	7	0.8	1036	12	US-10-145-877-142	Sequence 142, App
336	7	0.8	1036	12	US-10-145-958-142	Sequence 142, App
337	7	0.8	1036	12	US-10-146-787-142	Sequence 142, App
338	7	0.8	1036	12	US-10-146-790-142	Sequence 142, App
339	7	0.8	1036	12	US-10-146-793-142	Sequence 142, App
340	7	0.8	1036	12	US-10-147-480-142	Sequence 142, App
341	7	0.8	1036	12	US-10-147-485-142	Sequence 142, App
342	7	0.8	1036	12	US-10-147-486-142	Sequence 142, App
343	7	0.8	1036	12	US-10-147-487-142	Sequence 142, App
344	7	0.8	1036	12	US-10-147-490-142	Sequence 142, App
345	7	0.8	1036	12	US-10-147-494-142	Sequence 142, App
346	7	0.8	1036	12	US-10-147-498-142	Sequence 142, App
347	7	0.8	1036	12	US-10-147-514-142	Sequence 142, App
348	7	0.8	1036	12	US-10-147-524-142	Sequence 142, App
349	7	0.8	1036	12	US-10-152-379-142	Sequence 142, App
350	7	0.8	1036	12	US-10-152-394-142	Sequence 142, App
351	7	0.8	1036	12	US-10-152-406-142	Sequence 142, App
352	7	0.8	1036	12	US-10-156-847-142	Sequence 142, App
353	7	0.8	1036	12	US-10-157-778-142	Sequence 142, App
354	7	0.8	1036	12	US-10-157-799-142	Sequence 142, App
355	7	0.8	1036	12	US-10-160-504-142	Sequence 142, App
356	7	0.8	1036	12	US-10-145-634-142	Sequence 142, App
357	7	0.8	1036	12	US-10-147-520-142	Sequence 142, App
358	7	0.8	1036	12	US-10-157-781-142	Sequence 142, App
359	7	0.8	1036	12	US-10-176-989-142	Sequence 142, App
360	7	0.8	1036	12	US-10-147-491-142	Sequence 142, App
361	7	0.8	1036	12	US-10-152-378-142	Sequence 142, App
362	7	0.8	1036	12	US-10-152-382-142	Sequence 142, App
363	7	0.8	1036	12	US-10-152-383-142	Sequence 142, App
364	7	0.8	1036	12	US-10-152-384-142	Sequence 142, App
365	7	0.8	1036	12	US-10-152-387-142	Sequence 142, App
366	7	0.8	1036	12	US-10-152-389-142	Sequence 142, App
367	7	0.8	1036	12	US-10-152-390-142	Sequence 142, App
368	7	0.8	1036	12	US-10-152-392-142	Sequence 142, App
369	7	0.8	1036	12	US-10-153-756-142	Sequence 142, App
370	7	0.8	1036	12	US-10-157-779-142	Sequence 142, App
371	7	0.8	1036	12	US-10-157-784-142	Sequence 142, App
372	7	0.8	1036	12	US-10-157-797-142	Sequence 142, App
373	7	0.8	1036	12	US-10-158-491-142	Sequence 142, App
374	7	0.8	1036	12	US-10-139-980-142	Sequence 142, App
375	7	0.8	1036	12	US-10-142-762-142	Sequence 142, App
376	7	0.8	1036	12	US-10-142-764-142	Sequence 142, App
377	7	0.8	1036	12	US-10-142-766-142	Sequence 142, App
378	7	0.8	1036	12	US-10-145-625-142	Sequence 142, App
379	7	0.8	1036	12	US-10-145-627-142	Sequence 142, App
380	7	0.8	1036	12	US-10-145-960-142	Sequence 142, App

381	7	0.8	1036	12	US-10-145-962-142	Sequence 142, App
382	7	0.8	1036	12	US-10-146-789-142	Sequence 142, App
383	7	0.8	1036	12	US-10-147-483-142	Sequence 142, App
384	7	0.8	1036	12	US-10-147-496-142	Sequence 142, App
385	7	0.8	1036	12	US-10-147-505-142	Sequence 142, App
386	7	0.8	1036	12	US-10-147-516-142	Sequence 142, App
387	7	0.8	1036	12	US-10-152-398-142	Sequence 142, App
388	7	0.8	1036	12	US-10-145-750-142	Sequence 142, App
389	7	0.8	1036	12	US-10-152-373-142	Sequence 142, App
390	7	0.8	1036	12	US-10-121-044-142	Sequence 142, App
391	7	0.8	1036	12	US-10-121-055-142	Sequence 142, App
392	7	0.8	1036	12	US-10-121-057-142	Sequence 142, App
393	7	0.8	1036	12	US-10-121-058-142	Sequence 142, App
394	7	0.8	1036	12	US-10-121-059-142	Sequence 142, App
395	7	0.8	1036	12	US-10-121-060-142	Sequence 142, App
396	7	0.8	1036	12	US-10-123-109-142	Sequence 142, App
397	7	0.8	1036	12	US-10-123-154-142	Sequence 142, App
398	7	0.8	1036	12	US-10-123-157-142	Sequence 142, App
399	7	0.8	1036	12	US-10-123-906-142	Sequence 142, App
400	7	0.8	1036	12	US-10-124-814-142	Sequence 142, App
401	7	0.8	1036	12	US-10-124-816-142	Sequence 142, App
402	7	0.8	1036	12	US-10-124-820-142	Sequence 142, App
403	7	0.8	1036	12	US-10-125-704-142	Sequence 142, App
404	7	0.8	1036	12	US-10-125-927-142	Sequence 142, App
405	7	0.8	1036	12	US-10-121-046-142	Sequence 142, App
406	7	0.8	1036	12	US-10-123-156-142	Sequence 142, App
407	7	0.8	1036	12	US-10-123-214-142	Sequence 142, App
408	7	0.8	1036	12	US-10-125-805-142	Sequence 142, App
409	7	0.8	1036	12	US-10-142-889-142	Sequence 142, App
410	7	0.8	1036	12	US-10-145-874-142	Sequence 142, App
411	7	0.8	1036	12	US-10-147-497-142	Sequence 142, App
412	7	0.8	1036	12	US-10-152-371-142	Sequence 142, App
413	7	0.8	1036	12	US-10-152-374-142	Sequence 142, App
414	7	0.8	1036	12	US-10-152-375-142	Sequence 142, App
415	7	0.8	1036	12	US-10-152-377-142	Sequence 142, App
416	7	0.8	1036	12	US-10-152-386-142	Sequence 142, App
417	7	0.8	1036	12	US-10-152-391-142	Sequence 142, App
418	7	0.8	1036	12	US-10-152-399-142	Sequence 142, App
419	7	0.8	1036	12	US-10-156-848-142	Sequence 142, App
420	7	0.8	1036	12	US-10-157-785-142	Sequence 142, App
421	7	0.8	1036	12	US-10-157-794-142	Sequence 142, App
422	7	0.8	1036	12	US-10-157-796-142	Sequence 142, App
423	7	0.8	1036	12	US-10-160-500-142	Sequence 142, App
424	7	0.8	1036	12	US-10-121-048-142	Sequence 142, App
425	7	0.8	1036	12	US-10-121-052-142	Sequence 142, App
426	7	0.8	1036	12	US-10-121-053-142	Sequence 142, App
427	7	0.8	1036	12	US-10-121-054-142	Sequence 142, App
428	7	0.8	1036	12	US-10-121-063-142	Sequence 142, App
429	7	0.8	1036	12	US-10-123-212-142	Sequence 142, App
430	7	0.8	1036	12	US-10-123-213-142	Sequence 142, App
431	7	0.8	1036	12	US-10-123-291-142	Sequence 142, App
432	7	0.8	1036	12	US-10-123-322-142	Sequence 142, App
433	7	0.8	1036	12	US-10-123-771-142	Sequence 142, App
434	7	0.8	1036	12	US-10-123-911-142	Sequence 142, App
435	7	0.8	1036	12	US-10-124-821-142	Sequence 142, App
436	7	0.8	1036	12	US-10-124-823-142	Sequence 142, App
437	7	0.8	1036	12	US-10-125-931-142	Sequence 142, App
438	7	0.8	1036	12	US-10-125-932-142	Sequence 142, App
439	7	0.8	1036	12	US-10-152-385-142	Sequence 142, App
440	7	0.8	1036	12	US-10-152-393-142	Sequence 142, App
441	7	0.8	1036	12	US-10-152-396-142	Sequence 142, App
442	7	0.8	1036	12	US-10-153-552-142	Sequence 142, App
443	7	0.8	1036	12	US-10-153-840-142	Sequence 142, App
444	7	0.8	1036	12	US-10-156-841-142	Sequence 142, App
445	7	0.8	1036	12	US-10-156-842-142	Sequence 142, App
446	7	0.8	1036	12	US-10-156-844-142	Sequence 142, App
447	7	0.8	1036	12	US-10-156-845-142	Sequence 142, App
448	7	0.8	1036	12	US-10-156-846-142	Sequence 142, App
449	7	0.8	1036	12	US-10-123-913-142	Sequence 142, App
450	7	0.8	1036	12	US-10-139-963-142	Sequence 142, App
451	7	0.8	1036	12	US-10-140-020-142	Sequence 142, App
452	7	0.8	1036	12	US-10-140-023-142	Sequence 142, App
453	7	0.8	1036	12	US-10-140-473-142	Sequence 142, App
454	7	0.8	1036	12	US-10-140-806-142	Sequence 142, App
455	7	0.8	1036	12	US-10-140-809-142	Sequence 142, App
456	7	0.8	1036	12	US-10-140-810-142	Sequence 142, App
457	7	0.8	1036	12	US-10-140-863-142	Sequence 142, App
458	7	0.8	1036	12	US-10-140-865-142	Sequence 142, App
459	7	0.8	1036	12	US-10-141-699-142	Sequence 142, App
460	7	0.8	1036	12	US-10-141-701-142	Sequence 142, App
461	7	0.8	1036	12	US-10-141-703-142	Sequence 142, App
462	7	0.8	1036	12	US-10-141-706-142	Sequence 142, App
463	7	0.8	1036	12	US-10-141-754-142	Sequence 142, App
464	7	0.8	1036	12	US-10-141-757-142	Sequence 142, App
465	7	0.8	1036	12	US-10-141-760-142	Sequence 142, App
466	7	0.8	1036	12	US-10-141-762-142	Sequence 142, App
467	7	0.8	1036	12	US-10-142-425-142	Sequence 142, App
468	7	0.8	1036	12	US-10-142-428-142	Sequence 142, App
469	7	0.8	1036	12	US-10-142-429-142	Sequence 142, App
470	7	0.8	1036	12	US-10-142-430-142	Sequence 142, App
471	7	0.8	1036	12	US-10-142-884-142	Sequence 142, App
472	7	0.8	1036	12	US-10-143-027-142	Sequence 142, App
473	7	0.8	1036	12	US-10-143-113-142	Sequence 142, App
474	7	0.8	1036	12	US-10-143-115-142	Sequence 142, App
475	7	0.8	1036	12	US-10-144-956-142	Sequence 142, App
476	7	0.8	1036	12	US-10-144-958-142	Sequence 142, App
477	7	0.8	1036	12	US-10-145-632-142	Sequence 142, App
478	7	0.8	1036	12	US-10-145-749-142	Sequence 142, App
479	7	0.8	1036	12	US-10-145-753-142	Sequence 142, App
480	7	0.8	1036	12	US-10-145-871-142	Sequence 142, App
481	7	0.8	1036	12	US-10-145-878-142	Sequence 142, App
482	7	0.8	1036	12	US-10-146-730-142	Sequence 142, App
483	7	0.8	1036	12	US-10-146-792-142	Sequence 142, App
484	7	0.8	1036	12	US-10-146-794-142	Sequence 142, App
485	7	0.8	1036	12	US-10-147-489-142	Sequence 142, App
486	7	0.8	1036	12	US-10-147-507-142	Sequence 142, App
487	7	0.8	1036	12	US-10-147-535-142	Sequence 142, App
488	7	0.8	1036	12	US-10-147-537-142	Sequence 142, App
489	7	0.8	1036	12	US-10-152-376-142	Sequence 142, App
490	7	0.8	1036	12	US-10-152-381-142	Sequence 142, App
491	7	0.8	1036	12	US-10-152-400-142	Sequence 142, App
492	7	0.8	1036	12	US-10-153-585-142	Sequence 142, App
493	7	0.8	1036	12	US-10-156-843-142	Sequence 142, App
494	7	0.8	1036	12	US-10-157-780-142	Sequence 142, App
495	7	0.8	1036	12	US-10-157-786-142	Sequence 142, App
496	7	0.8	1036	12	US-10-157-800-142	Sequence 142, App
497	7	0.8	1036	12	US-10-157-801-142	Sequence 142, App
498	7	0.8	1036	12	US-10-157-802-142	Sequence 142, App
499	7	0.8	1036	12	US-10-158-784-142	Sequence 142, App
500	7	0.8	1036	12	US-10-158-789-142	Sequence 142, App
501	7	0.8	1036	12	US-10-158-791-142	Sequence 142, App
502	7	0.8	1036	12	US-10-192-011-142	Sequence 142, App
503	7	0.8	1036	12	US-10-147-528-142	Sequence 142, App
504	7	0.8	1036	12	US-10-236-031B-32	Sequence 32, Appl
505	7	0.8	1036	12	US-10-128-692A-142	Sequence 142, App
506	7	0.8	1036	12	US-10-140-927-142	Sequence 142, App
507	7	0.8	1036	15	US-10-028-072-142	Sequence 142, App
508	7	0.8	1036	15	US-10-121-049-142	Sequence 142, App
509	7	0.8	1036	15	US-10-123-904-142	Sequence 142, App
510	7	0.8	1036	15	US-10-140-470-142	Sequence 142, App
511	7	0.8	1036	15	US-10-175-746-142	Sequence 142, App
512	7	0.8	1036	15	US-10-176-918-142	Sequence 142, App
513	7	0.8	1036	15	US-10-176-921-142	Sequence 142, App
514	7	0.8	1036	15	US-10-137-865-142	Sequence 142, App
515	7	0.8	1036	15	US-10-140-474-142	Sequence 142, App
516	7	0.8	1036	15	US-10-142-431-142	Sequence 142, App
517	7	0.8	1036	15	US-10-143-114-142	Sequence 142, App
518	7	0.8	1036	15	US-10-140-002-142	Sequence 142, App
519	7	0.8	1036	15	US-10-142-419-142	Sequence 142, App
520	7	0.8	1036	15	US-10-123-262-142	Sequence 142, App
521	7	0.8	1036	15	US-10-142-423-142	Sequence 142, App
522	7	0.8	1036	15	US-10-121-050-142	Sequence 142, App
523	7	0.8	1036	15	US-10-141-755-142	Sequence 142, App
524	7	0.8	1036	15	US-10-143-032-142	Sequence 142, App
525	7	0.8	1036	15	US-10-123-108-142	Sequence 142, App
526	7	0.8	1036	15	US-10-123-236-142	Sequence 142, App

527	7	0.8	1036	15	US-10-123-261-142	Sequence 142, App
528	7	0.8	1036	15	US-10-140-921-142	Sequence 142, App
529	7	0.8	1036	15	US-10-140-928-142	Sequence 142, App
530	7	0.8	1036	15	US-10-121-045-142	Sequence 142, App
531	7	0.8	1036	15	US-10-123-292-142	Sequence 142, App
532	7	0.8	1036	15	US-10-123-903-142	Sequence 142, App
533	7	0.8	1036	15	US-10-124-819-142	Sequence 142, App
534	7	0.8	1036	15	US-10-124-822-142	Sequence 142, App
535	7	0.8	1036	15	US-10-140-925-142	Sequence 142, App
536	7	0.8	1036	15	US-10-160-498-142	Sequence 142, App
537	7	0.8	1036	15	US-10-124-824-142	Sequence 142, App
538	7	0.8	1036	15	US-10-127-825A-142	Sequence 142, App
539	7	0.8	1036	15	US-10-127-829A-142	Sequence 142, App
540	7	0.8	1036	15	US-10-127-835A-142	Sequence 142, App
541	7	0.8	1036	15	US-10-127-839A-142	Sequence 142, App
542	7	0.8	1036	15	US-10-127-901A-142	Sequence 142, App
543	7	0.8	1036	15	US-10-128-693A-142	Sequence 142, App
544	7	0.8	1036	15	US-10-131-813A-142	Sequence 142, App
545	7	0.8	1036	15	US-10-131-818A-142	Sequence 142, App
546	7	0.8	1036	15	US-10-131-823A-142	Sequence 142, App
547	7	0.8	1036	15	US-10-131-824A-142	Sequence 142, App
548	7	0.8	1036	15	US-10-131-830A-142	Sequence 142, App
549	7	0.8	1036	15	US-10-131-837A-142	Sequence 142, App
550	7	0.8	1036	15	US-10-137-872A-142	Sequence 142, App
551	7	0.8	1036	15	US-10-147-500-142	Sequence 142, App
552	7	0.8	1036	15	US-10-147-502-142	Sequence 142, App
553	7	0.8	1036	15	US-10-147-515-142	Sequence 142, App
554	7	0.8	1036	15	US-10-147-517-142	Sequence 142, App
555	7	0.8	1036	15	US-10-147-526-142	Sequence 142, App
556	7	0.8	1036	15	US-10-147-527-142	Sequence 142, App
557	7	0.8	1036	15	US-10-121-041-142	Sequence 142, App
558	7	0.8	1036	15	US-10-121-043-142	Sequence 142, App
559	7	0.8	1036	15	US-10-121-047-142	Sequence 142, App
560	7	0.8	1036	15	US-10-123-215-142	Sequence 142, App
561	7	0.8	1036	15	US-10-123-902-142	Sequence 142, App
562	7	0.8	1036	15	US-10-123-908-142	Sequence 142, App
563	7	0.8	1036	15	US-10-123-909-142	Sequence 142, App
564	7	0.8	1036	15	US-10-123-910-142	Sequence 142, App
565	7	0.8	1036	15	US-10-124-813-142	Sequence 142, App
566	7	0.8	1036	15	US-10-124-817-142	Sequence 142, App
567	7	0.8	1036	15	US-10-125-922-142	Sequence 142, App
568	7	0.8	1036	15	US-10-125-924-142	Sequence 142, App
569	7	0.8	1036	15	US-10-140-860-142	Sequence 142, App
570	7	0.8	1036	15	US-10-142-417-142	Sequence 142, App
571	7	0.8	1036	15	US-10-147-519-142	Sequence 142, App
572	7	0.8	1036	15	US-10-157-782-142	Sequence 142, App
573	7	0.8	1036	15	US-10-152-395-142	Sequence 142, App
574	7	0.8	1036	15	US-10-125-926A-142	Sequence 142, App
575	7	0.8	1036	15	US-10-125-930A-142	Sequence 142, App
576	7	0.8	1036	15	US-10-127-831A-142	Sequence 142, App
577	7	0.8	1036	15	US-10-127-837A-142	Sequence 142, App
578	7	0.8	1036	15	US-10-127-838B-142	Sequence 142, App
579	7	0.8	1036	15	US-10-127-842A-142	Sequence 142, App
580	7	0.8	1036	15	US-10-127-843A-142	Sequence 142, App
581	7	0.8	1036	15	US-10-127-845A-142	Sequence 142, App
582	7	0.8	1036	15	US-10-127-846A-142	Sequence 142, App
583	7	0.8	1036	15	US-10-127-848A-142	Sequence 142, App
584	7	0.8	1036	15	US-10-127-849A-142	Sequence 142, App
585	7	0.8	1036	15	US-10-127-850A-142	Sequence 142, App
586	7	0.8	1036	15	US-10-127-851A-142	Sequence 142, App
587	7	0.8	1036	15	US-10-128-684A-142	Sequence 142, App
588	7	0.8	1036	15	US-10-128-686A-142	Sequence 142, App
589	7	0.8	1036	15	US-10-128-690A-142	Sequence 142, App
590	7	0.8	1036	15	US-10-128-691A-142	Sequence 142, App
591	7	0.8	1036	15	US-10-131-819A-142	Sequence 142, App
592	7	0.8	1036	15	US-10-131-829A-142	Sequence 142, App
593	7	0.8	1036	15	US-10-131-836A-142	Sequence 142, App
594	7	0.8	1036	15	US-10-146-729-142	Sequence 142, App
595	7	0.8	1036	15	US-10-146-791-142	Sequence 142, App
596	7	0.8	1036	15	US-10-147-484-142	Sequence 142, App
597	7	0.8	1036	15	US-10-147-508-142	Sequence 142, App
598	7	0.8	1036	15	US-10-147-512-142	Sequence 142, App
599	7	0.8	1036	15	US-10-152-724A-2	Sequence 2, Appli
600	7	0.8	1036	15	US-10-175-735-142	Sequence 142, App
601	7	0.8	1036	15	US-10-121-040-142	Sequence 142, App
602	7	0.8	1036	15	US-10-121-056-142	Sequence 142, App
603	7	0.8	1036	15	US-10-121-061-142	Sequence 142, App
604	7	0.8	1036	15	US-10-123-235-142	Sequence 142, App
605	7	0.8	1036	15	US-10-124-818-142	Sequence 142, App
606	7	0.8	1036	15	US-10-137-868-142	Sequence 142, App
607	7	0.8	1036	15	US-10-147-492-142	Sequence 142, App
608	7	0.8	1036	15	US-10-158-782-142	Sequence 142, App
609	7	0.8	1036	15	US-10-123-905-142	Sequence 142, App
610	7	0.8	1036	15	US-10-123-907-142	Sequence 142, App
611	7	0.8	1036	15	US-10-124-815-142	Sequence 142, App
612	7	0.8	1036	15	US-10-125-921A-142	Sequence 142, App
613	7	0.8	1036	15	US-10-125-928A-142	Sequence 142, App
614	7	0.8	1036	15	US-10-127-821A-142	Sequence 142, App
615	7	0.8	1036	15	US-10-127-822A-142	Sequence 142, App
616	7	0.8	1036	15	US-10-127-824A-142	Sequence 142, App
617	7	0.8	1036	15	US-10-127-826A-142	Sequence 142, App
618	7	0.8	1036	15	US-10-127-827A-142	Sequence 142, App
619	7	0.8	1036	15	US-10-127-828A-142	Sequence 142, App
620	7	0.8	1036	15	US-10-127-830A-142	Sequence 142, App
621	7	0.8	1036	15	US-10-127-832A-142	Sequence 142, App
622	7	0.8	1036	15	US-10-127-833A-142	Sequence 142, App
623	7	0.8	1036	15	US-10-127-834A-142	Sequence 142, App
624	7	0.8	1036	15	US-10-127-836A-142	Sequence 142, App
625	7	0.8	1036	15	US-10-127-841A-142	Sequence 142, App
626	7	0.8	1036	15	US-10-127-844A-142	Sequence 142, App
627	7	0.8	1036	15	US-10-128-687A-142	Sequence 142, App
628	7	0.8	1036	15	US-10-128-688A-142	Sequence 142, App
629	7	0.8	1036	15	US-10-128-689A-142	Sequence 142, App
630	7	0.8	1036	15	US-10-128-694A-142	Sequence 142, App
631	7	0.8	1036	15	US-10-131-825A-142	Sequence 142, App
632	7	0.8	1036	15	US-10-230-417-142	Sequence 142, App
633	7	0.8	1036	15	US-10-131-815A-142	Sequence 142, App
634	7	0.8	1036	15	US-10-131-817A-142	Sequence 142, App
635	7	0.8	1036	15	US-10-131-821A-142	Sequence 142, App
636	7	0.8	1036	15	US-10-131-822A-142	Sequence 142, App
637	7	0.8	1036	15	US-10-131-828A-142	Sequence 142, App
638	7	0.8	1036	15	US-10-131-835A-142	Sequence 142, App
639	7	0.8	1036	15	US-10-137-864A-142	Sequence 142, App
640	7	0.8	1036	15	US-10-137-869A-142	Sequence 142, App
641	7	0.8	1036	15	US-10-147-523-142	Sequence 142, App
642	7	0.8	1036	15	US-10-158-785-142	Sequence 142, App
643	7	0.8	1036	15	US-10-121-051-142	Sequence 142, App
644	7	0.8	1036	15	US-10-121-042-142	Sequence 142, App
645	7	0.8	1036	15	US-10-123-912-142	Sequence 142, App
646	7	0.8	1036	15	US-10-192-007-142	Sequence 142, App
647	7	0.8	1036	15	US-10-194-359-142	Sequence 142, App
648	7	0.8	1036	15	US-10-127-847A-142	Sequence 142, App
649	7	0.8	1036	15	US-10-175-590-142	Sequence 142, App
650	7	0.8	1036	16	US-10-137-866-142	Sequence 142, App
651	7	0.8	1036	16	US-10-146-726-142	Sequence 142, App
652	7	0.8	1036	16	US-10-146-727-142	Sequence 142, App
653	7	0.8	1036	16	US-10-146-788-142	Sequence 142, App
654	7	0.8	1036	16	US-10-152-380-142	Sequence 142, App
655	7	0.8	1036	16	US-10-153-934-142	Sequence 142, App
656	7	0.8	1040	9	US-09-864-761-38325	Sequence 38325, A
657	7	0.8	1041	10	US-09-978-295A-498	Sequence 498, App
658	7	0.8	1041	10	US-09-978-697-498	Sequence 498, App
659	7	0.8	1041	10	US-09-978-192A-498	Sequence 498, App
660	7	0.8	1041	10	US-09-999-832A-498	Sequence 498, App
661	7	0.8	1041	11	US-09-978-189-498	Sequence 498, App
662	7	0.8	1041	11	US-09-978-608A-498	Sequence 498, App
663	7	0.8	1041	11	US-09-978-585A-498	Sequence 498, App
664	7	0.8	1041	11	US-09-978-191A-498	Sequence 498, App
665	7	0.8	1041	11	US-09-978-403A-498	Sequence 498, App
666	7	0.8	1041	11	US-09-978-564A-498	Sequence 498, App
667	7	0.8	1041	11	US-09-999-833A-498	Sequence 498, App
668	7	0.8	1041	11	US-09-981-915A-498	Sequence 498, App
669	7	0.8	1041	11	US-09-978-824-498	Sequence 498, App
670	7	0.8	1041	11	US-09-918-585A-498	Sequence 498, App
671	7	0.8	1041	11	US-09-978-423A-498	Sequence 498, App
672	7	0.8	1041	11	US-09-978-193A-498	Sequence 498, App

673	7	0.8	1041	11	US-09-999-830A-498	Sequence 498, App
674	7	0.8	1041	11	US-09-978-757A-498	Sequence 498, App
675	7	0.8	1041	11	US-09-978-187B-498	Sequence 498, App
676	7	0.8	1041	11	US-09-954-987B-184	Sequence 184, App
677	7	0.8	1041	11	US-09-954-987B-186	Sequence 186, App
678	7	0.8	1041	11	US-09-978-643A-498	Sequence 498, App
679	7	0.8	1041	12	US-09-978-375A-498	Sequence 498, App
680	7	0.8	1041	12	US-09-978-188A-498	Sequence 498, App
681	7	0.8	1041	12	US-09-978-298A-498	Sequence 498, App
682	7	0.8	1041	12	US-10-143-031A-498	Sequence 498, App
683	7	0.8	1041	12	US-10-272-502A-26	Sequence 26, Appl
684	7	0.8	1041	12	US-10-002-967A-498	Sequence 498, App
685	7	0.8	1041	12	US-10-017-083A-498	Sequence 498, App
686	7	0.8	1041	12	US-10-143-030A-498	Sequence 498, App
687	7	0.8	1041	12	US-10-145-128A-498	Sequence 498, App
688	7	0.8	1041	12	US-10-017-191A-498	Sequence 498, App
689	7	0.8	1041	12	US-10-143-028A-498	Sequence 498, App
690	7	0.8	1041	12	US-10-143-029A-498	Sequence 498, App
691	7	0.8	1041	12	US-10-145-089A-498	Sequence 498, App
692	7	0.8	1041	12	US-10-013-926A-498	Sequence 498, App
693	7	0.8	1041	12	US-10-145-017A-498	Sequence 498, App
694	7	0.8	1041	12	US-10-164-728A-498	Sequence 498, App
695	7	0.8	1041	12	US-10-165-067A-498	Sequence 498, App
696	7	0.8	1041	12	US-10-145-124A-498	Sequence 498, App
697	7	0.8	1041	12	US-10-160-502A-498	Sequence 498, App
698	7	0.8	1041	12	US-10-165-247A-498	Sequence 498, App
699	7	0.8	1041	12	US-09-978-194A-498	Sequence 498, App
700	7	0.8	1041	12	US-09-978-681A-498	Sequence 498, App
701	7	0.8	1041	12	US-09-999-829A-498	Sequence 498, App
702	7	0.8	1041	12	US-10-013-922A-498	Sequence 498, App
703	7	0.8	1041	12	US-10-017-086A-498	Sequence 498, App
704	7	0.8	1041	12	US-10-145-087A-498	Sequence 498, App
705	7	0.8	1041	12	US-10-164-829A-498	Sequence 498, App
706	7	0.8	1041	12	US-10-164-929A-498	Sequence 498, App
707	7	0.8	1041	12	US-09-978-299A-498	Sequence 498, App
708	7	0.8	1041	12	US-09-978-544A-498	Sequence 498, App
709	7	0.8	1041	12	US-09-978-665A-498	Sequence 498, App
710	7	0.8	1041	12	US-09-978-802A-498	Sequence 498, App
711	7	0.8	1041	12	US-10-013-924A-498	Sequence 498, App
712	7	0.8	1041	12	US-10-020-445A-498	Sequence 498, App
713	7	0.8	1041	12	US-10-017-084A-498	Sequence 498, App
714	7	0.8	1041	12	US-10-017-085A-498	Sequence 498, App
715	7	0.8	1041	12	US-10-013-916A-498	Sequence 498, App
716	7	0.8	1041	12	US-10-143-026B-498	Sequence 498, App
717	7	0.8	1041	12	US-10-013-918A-498	Sequence 498, App
718	7	0.8	1041	12	US-10-013-923A-498	Sequence 498, App
719	7	0.8	1041	12	US-10-013-925A-498	Sequence 498, App
720	7	0.8	1041	12	US-10-013-927A-498	Sequence 498, App
721	7	0.8	1041	12	US-10-013-928A-498	Sequence 498, App
722	7	0.8	1041	12	US-10-162-522A-498	Sequence 498, App
723	7	0.8	1041	12	US-10-407-952-28	Sequence 28, Appl
724	7	0.8	1041	12	US-10-013-919A-498	Sequence 498, App
725	7	0.8	1041	12	US-10-013-920A-498	Sequence 498, App
726	7	0.8	1041	12	US-10-145-093A-498	Sequence 498, App
727	7	0.8	1041	15	US-10-095-627-3	Sequence 3, Appli
728	7	0.8	1041	15	US-10-017-081A-498	Sequence 498, App
729	7	0.8	1041	15	US-10-167-749-498	Sequence 498, App
730	7	0.8	1041	15	US-10-013-921A-498	Sequence 498, App
731	7	0.8	1041	15	US-10-013-929A-498	Sequence 498, App
732	7	0.8	1041	15	US-10-016-177A-498	Sequence 498, App
733	7	0.8	1041	15	US-10-235-767-3	Sequence 3, Appli
734	7	0.8	1041	15	US-10-166-709A-498	Sequence 498, App
735	7	0.8	1059	11	US-09-954-987B-187	Sequence 187, App
736	7	0.8	1059	12	US-10-407-952-30	Sequence 30, Appl
737	7	0.8	1109	10	US-09-529-063-55	Sequence 55, Appl
738	7	0.8	1109	12	US-10-414-378-55	Sequence 55, Appl
739	7	0.8	1140	12	US-10-369-493-18194	Sequence 18194, A
740	7	0.8	1163	12	US-10-452-024-107	Sequence 107, App
741	7	0.8	1211	12	US-10-017-161-824	Sequence 824, App
742	7	0.8	1247	10	US-09-738-626-4751	Sequence 4751, Ap
743	7	0.8	1266	12	US-10-320-797-3352	Sequence 3352, Ap
744	7	0.8	1275	15	US-10-128-714-3125	Sequence 3125, Ap
745	7	0.8	1280	12	US-10-452-024-162	Sequence 162, App
746	7	0.8	1290	10	US-09-881-752A-138	Sequence 138, App
747	7	0.8	1290	12	US-10-360-101-220	Sequence 220, App
748	7	0.8	1290	12	US-10-452-024-119	Sequence 119, App
749	7	0.8	1291	12	US-10-354-774-42	Sequence 42, Appl
750	7	0.8	1291	12	US-10-271-012-42	Sequence 42, Appl
751	7	0.8	1291	12	US-10-452-024-2	Sequence 2, Appli
752	7	0.8	1291	12	US-10-452-024-118	Sequence 118, App
753	7	0.8	1291	12	US-10-452-024-121	Sequence 121, App
754	7	0.8	1291	12	US-10-452-024-122	Sequence 122, App
755	7	0.8	1291	12	US-10-452-024-123	Sequence 123, App
756	7	0.8	1291	12	US-10-452-024-142	Sequence 142, App
757	7	0.8	1291	15	US-10-156-761-14161	Sequence 14161, A
758	7	0.8	1435	15	US-10-128-714-8125	Sequence 8125, Ap
759	7	0.8	1908	15	US-10-128-714-3475	Sequence 3475, Ap
760	7	0.8	2022	15	US-10-128-714-8475	Sequence 8475, Ap
761	7	0.8	2099	15	US-10-128-714-3290	Sequence 3290, Ap
762	7	0.8	2405	15	US-10-128-714-8290	Sequence 8290, Ap
763	7	0.8	2427	12	US-10-369-493-6734	Sequence 6734, Ap
764	7	0.8	2630	12	US-10-369-493-3693	Sequence 3693, Ap
765	7	0.8	3234	12	US-10-093-463-168	Sequence 168, App
766	7	0.8	3390	12	US-10-369-493-21101	Sequence 21101, A
767	7	0.8	3430	12	US-10-224-999A-3473	Sequence 3473, Ap
768	7	0.8	4999	10	US-09-976-059-14	Sequence 14, Appl
769	7	0.8	19695	12	US-10-084-846A-3	Sequence 3, Appli
770	6	0.6	8	12	US-10-351-641-1592	Sequence 1592, Ap
771	6	0.6	9	12	US-10-034-974-56	Sequence 56, Appl
772	6	0.6	9	12	US-09-793-451-530	Sequence 530, App
773	6	0.6	9	12	US-09-793-451-642	Sequence 642, App
774	6	0.6	9	12	US-10-283-722-530	Sequence 530, App
775	6	0.6	9	12	US-10-283-722-642	Sequence 642, App
776	6	0.6	9	12	US-10-283-903-530	Sequence 530, App
777	6	0.6	9	12	US-10-283-903-642	Sequence 642, App
778	6	0.6	10	7	US-08-344-824-61	Sequence 61, Appl
779	6	0.6	10	11	US-09-572-404B-1963	Sequence 1963, Ap
780	6	0.6	10	12	US-09-572-270A-1057	Sequence 1057, Ap
781	6	0.6	10	12	US-09-793-451-77	Sequence 77, Appl
782	6	0.6	10	12	US-09-793-451-81	Sequence 81, Appl
783	6	0.6	10	12	US-09-793-451-412	Sequence 412, App
784	6	0.6	10	12	US-10-283-722-77	Sequence 77, Appl
785	6	0.6	10	12	US-10-283-722-81	Sequence 81, Appl
786	6	0.6	10	12	US-10-283-722-412	Sequence 412, App
787	6	0.6	10	12	US-10-117-937-433	Sequence 433, App
788	6	0.6	10	12	US-10-283-903-77	Sequence 77, Appl
789	6	0.6	10	12	US-10-283-903-81	Sequence 81, Appl
790	6	0.6	10	12	US-10-283-903-412	Sequence 412, App
791	6	0.6	10	15	US-10-094-699-88	Sequence 88, Appl
792	6	0.6	10	15	US-10-279-991-9	Sequence 9, Appli
793	6	0.6	12	12	US-09-898-860-30	Sequence 30, Appl
794	6	0.6	12	12	US-10-366-125-7	Sequence 7, Appli
795	6	0.6	13	14	US-10-015-536-19	Sequence 19, Appl
796	6	0.6	13	14	US-10-015-536-21	Sequence 21, Appl
797	6	0.6	15	12	US-10-034-974-40	Sequence 40, Appl
798	6	0.6	15	15	US-10-279-991-6	Sequence 6, Appli
799	6	0.6	16	9	US-09-770-940-14	Sequence 14, Appl
800	6	0.6	16	12	US-10-165-528-14	Sequence 14, Appl
801	6	0.6	17	12	US-10-280-066-235	Sequence 235, App
802	6	0.6	17	12	US-10-425-328-22	Sequence 22, Appl
803	6	0.6	18	12	US-10-145-206-130	Sequence 130, App
804	6	0.6	18	12	US-10-408-166-454	Sequence 454, App
805	6	0.6	19	12	US-10-408-166-453	Sequence 453, App
806	6	0.6	20	11	US-09-858-935B-89	Sequence 89, Appl
807	6	0.6	20	12	US-10-408-166-452	Sequence 452, App
808	6	0.6	20	15	US-10-044-708A-10	Sequence 10, Appl
809	6	0.6	21	12	US-10-408-166-451	Sequence 451, App
810	6	0.6	22	12	US-10-408-166-450	Sequence 450, App
811	6	0.6	23	12	US-10-408-166-449	Sequence 449, App
812	6	0.6	24	9	US-09-789-404-16	Sequence 16, Appl
813	6	0.6	24	9	US-09-875-494-1	Sequence 1, Appli
814	6	0.6	24	9	US-09-875-494-2	Sequence 2, Appli
815	6	0.6	24	9	US-09-875-494-23	Sequence 23, Appl
816	6	0.6	24	12	US-09-896-095-251	Sequence 251, App
817	6	0.6	24	12	US-10-273-973-9	Sequence 9, Appli
818	6	0.6	24	12	US-10-408-166-448	Sequence 448, App

819	6	0.6	25	12	US-10-408-166-447	Sequence 447, App
820	6	0.6	26	9	US-09-864-761-45335	Sequence 45335, A
821	6	0.6	26	12	US-10-408-166-446	Sequence 446, App
822	6	0.6	26	15	US-10-012-542-383	Sequence 383, App
823	6	0.6	27	12	US-10-408-166-445	Sequence 445, App
824	6	0.6	28	9	US-09-770-940-15	Sequence 15, Appl
825	6	0.6	28	9	US-09-864-761-34971	Sequence 34971, A
826	6	0.6	28	9	US-09-864-761-41533	Sequence 41533, A
827	6	0.6	28	12	US-10-165-528-15	Sequence 15, Appl
828	6	0.6	28	12	US-10-408-166-444	Sequence 444, App
829	6	0.6	29	12	US-10-029-386-31377	Sequence 31377, A
830	6	0.6	29	12	US-10-408-166-443	Sequence 443, App
831	6	0.6	30	9	US-09-789-561-157	Sequence 157, App
832	6	0.6	30	10	US-09-790-622-6	Sequence 6, Appli
833	6	0.6	30	12	US-10-408-166-442	Sequence 442, App
834	6	0.6	30	15	US-10-141-953-6	Sequence 6, Appli
835	6	0.6	31	9	US-09-864-761-35950	Sequence 35950, A
836	6	0.6	31	12	US-10-408-166-441	Sequence 441, App
837	6	0.6	32	12	US-10-408-166-440	Sequence 440, App
838	6	0.6	33	9	US-09-864-761-35444	Sequence 35444, A
839	6	0.6	33	12	US-10-408-166-439	Sequence 439, App
840	6	0.6	33	14	US-10-038-045-14	Sequence 14, Appl
841	6	0.6	34	12	US-10-408-166-438	Sequence 438, App
842	6	0.6	35	12	US-10-408-166-437	Sequence 437, App
843	6	0.6	35	14	US-10-029-217A-23	Sequence 23, Appl
844	6	0.6	36	9	US-09-864-761-44661	Sequence 44661, A
845	6	0.6	36	12	US-10-408-166-436	Sequence 436, App
846	6	0.6	37	9	US-09-864-761-43447	Sequence 43447, A
847	6	0.6	37	11	US-09-776-724A-88	Sequence 88, Appl
848	6	0.6	37	12	US-10-029-386-33604	Sequence 33604, A
849	6	0.6	37	12	US-10-408-166-435	Sequence 435, App
850	6	0.6	38	10	US-09-925-300-1127	Sequence 1127, Ap
851	6	0.6	38	12	US-10-408-166-434	Sequence 434, App
852	6	0.6	39	12	US-10-408-166-385	Sequence 385, App
853	6	0.6	39	12	US-10-408-166-433	Sequence 433, App
854	6	0.6	40	12	US-10-408-166-384	Sequence 384, App
855	6	0.6	40	12	US-10-408-166-432	Sequence 432, App
856	6	0.6	41	12	US-10-408-166-383	Sequence 383, App
857	6	0.6	41	12	US-10-408-166-431	Sequence 431, App
858	6	0.6	42	12	US-10-408-166-382	Sequence 382, App
859	6	0.6	42	12	US-10-408-166-430	Sequence 430, App
860	6	0.6	43	9	US-09-864-761-43508	Sequence 43508, A
861	6	0.6	43	12	US-10-321-857-108	Sequence 108, App
862	6	0.6	43	12	US-10-318-675-108	Sequence 108, App
863	6	0.6	43	12	US-10-408-166-381	Sequence 381, App
864	6	0.6	43	12	US-10-408-166-429	Sequence 429, App
865	6	0.6	44	11	US-09-895-298-126	Sequence 126, App
866	6	0.6	44	12	US-10-195-730-339	Sequence 339, App
867	6	0.6	44	12	US-10-408-166-380	Sequence 380, App
868	6	0.6	44	12	US-10-408-166-428	Sequence 428, App
869	6	0.6	45	9	US-09-864-761-39649	Sequence 39649, A
870	6	0.6	45	9	US-09-864-761-48064	Sequence 48064, A
871	6	0.6	45	12	US-10-408-166-379	Sequence 379, App
872	6	0.6	45	12	US-10-408-166-427	Sequence 427, App
873	6	0.6	46	9	US-09-205-658-83	Sequence 83, Appl
874	6	0.6	46	9	US-09-844-353A-83	Sequence 83, Appl
875	6	0.6	46	9	US-09-864-761-35341	Sequence 35341, A
876	6	0.6	46	12	US-09-963-693-83	Sequence 83, Appl
877	6	0.6	46	12	US-10-408-166-378	Sequence 378, App
878	6	0.6	46	12	US-10-408-166-426	Sequence 426, App
879	6	0.6	47	12	US-10-029-386-31159	Sequence 31159, A
880	6	0.6	47	12	US-10-408-166-377	Sequence 377, App
881	6	0.6	47	12	US-10-408-166-425	Sequence 425, App
882	6	0.6	48	12	US-09-933-767-358	Sequence 358, App
883	6	0.6	48	12	US-10-408-166-376	Sequence 376, App
884	6	0.6	48	12	US-10-408-166-424	Sequence 424, App
885	6	0.6	48	15	US-10-023-282-358	Sequence 358, App
886	6	0.6	49	9	US-09-789-561-127	Sequence 127, App
887	6	0.6	49	12	US-09-833-245-2102	Sequence 2102, Ap
888	6	0.6	49	12	US-10-408-166-375	Sequence 375, App
889	6	0.6	49	12	US-10-408-166-423	Sequence 423, App
890	6	0.6	50	10	US-09-911-826A-15	Sequence 15, Appl
891	6	0.6	50	12	US-10-190-435-197	Sequence 197, App
892	6	0.6	50	12	US-10-408-166-374	Sequence 374, App
893	6	0.6	50	12	US-10-408-166-422	Sequence 422, App
894	6	0.6	50	15	US-10-106-698-8298	Sequence 8298, Ap
895	6	0.6	51	10	US-09-989-903-34	Sequence 34, Appl
896	6	0.6	51	12	US-10-408-166-373	Sequence 373, App
897	6	0.6	51	15	US-10-068-564-34	Sequence 34, Appl
898	6	0.6	52	9	US-09-864-761-42840	Sequence 42840, A
899	6	0.6	53	12	US-10-105-232-370	Sequence 370, App
900	6	0.6	53	12	US-10-189-437-357	Sequence 357, App
901	6	0.6	53	12	US-10-408-166-189	Sequence 189, App
902	6	0.6	53	15	US-10-106-698-5717	Sequence 5717, Ap
903	6	0.6	54	12	US-09-864-408A-7516	Sequence 7516, Ap
904	6	0.6	54	12	US-10-408-166-190	Sequence 190, App
905	6	0.6	55	9	US-09-764-878-136	Sequence 136, App
906	6	0.6	55	9	US-09-764-860-473	Sequence 473, App
907	6	0.6	55	12	US-10-212-872-473	Sequence 473, App
908	6	0.6	55	12	US-10-408-166-191	Sequence 191, App
909	6	0.6	55	15	US-10-079-854-136	Sequence 136, App
910	6	0.6	55	15	US-10-074-095-473	Sequence 473, App
911	6	0.6	56	12	US-10-408-166-192	Sequence 192, App
912	6	0.6	57	12	US-10-029-386-30158	Sequence 30158, A
913	6	0.6	57	12	US-10-408-166-193	Sequence 193, App
914	6	0.6	58	10	US-09-948-080-22	Sequence 22, Appl
915	6	0.6	58	12	US-10-408-166-194	Sequence 194, App
916	6	0.6	58	15	US-10-156-761-8226	Sequence 8226, Ap
917	6	0.6	59	11	US-09-892-877-339	Sequence 339, App
918	6	0.6	59	11	US-09-948-783-352	Sequence 352, App
919	6	0.6	59	12	US-10-408-166-195	Sequence 195, App
920	6	0.6	60	12	US-10-408-166-196	Sequence 196, App
921	6	0.6	60	15	US-10-083-357-817	Sequence 817, App
922	6	0.6	61	9	US-09-764-869-1160	Sequence 1160, Ap
923	6	0.6	61	12	US-10-227-577-1160	Sequence 1160, Ap
924	6	0.6	61	12	US-10-408-166-197	Sequence 197, App
925	6	0.6	61	15	US-10-091-504-1160	Sequence 1160, Ap
926	6	0.6	62	11	US-09-764-891-3703	Sequence 3703, Ap
927	6	0.6	62	12	US-09-864-408A-5028	Sequence 5028, Ap
928	6	0.6	62	12	US-10-408-166-198	Sequence 198, App
929	6	0.6	63	10	US-09-738-626-3628	Sequence 3628, Ap
930	6	0.6	63	12	US-10-408-166-199	Sequence 199, App
931	6	0.6	64	12	US-09-933-767-570	Sequence 570, App
932	6	0.6	64	12	US-10-408-166-200	Sequence 200, App
933	6	0.6	64	15	US-10-023-282-570	Sequence 570, App
934	6	0.6	65	12	US-10-029-386-33591	Sequence 33591, A
935	6	0.6	65	12	US-10-408-166-201	Sequence 201, App
936	6	0.6	66	12	US-09-864-408A-5204	Sequence 5204, Ap
937	6	0.6	66	12	US-10-408-166-202	Sequence 202, App
938	6	0.6	67	12	US-09-864-408A-186	Sequence 186, App
939	6	0.6	67	12	US-10-408-166-203	Sequence 203, App
940	6	0.6	68	11	US-09-823-187-47	Sequence 47, Appl
941	6	0.6	68	12	US-10-408-166-204	Sequence 204, App
942	6	0.6	69	9	US-09-864-761-44802	Sequence 44802, A
943	6	0.6	69	10	US-09-738-626-5569	Sequence 5569, Ap
944	6	0.6	69	12	US-10-029-386-27796	Sequence 27796, A
945	6	0.6	69	12	US-10-408-166-205	Sequence 205, App
946	6	0.6	70	11	US-09-764-891-3021	Sequence 3021, Ap
947	6	0.6	70	12	US-10-408-166-206	Sequence 206, App
948	6	0.6	71	10	US-09-943-123-15	Sequence 15, Appl
949	6	0.6	71	12	US-10-408-166-207	Sequence 207, App
950	6	0.6	71	15	US-10-050-704-112	Sequence 112, App
951	6	0.6	72	11	US-09-764-891-2973	Sequence 2973, Ap
952	6	0.6	72	12	US-10-266-829-103	Sequence 103, App
953	6	0.6	72	12	US-10-408-166-208	Sequence 208, App
954	6	0.6	73	12	US-10-408-166-209	Sequence 209, App
955	6	0.6	74	11	US-09-468-147-190	Sequence 190, App
956	6	0.6	74	11	US-09-468-147-197	Sequence 197, App
957	6	0.6	74	12	US-10-029-386-28264	Sequence 28264, A
958	6	0.6	74	12	US-09-864-408A-6794	Sequence 6794, Ap
959	6	0.6	74	12	US-10-408-166-210	Sequence 210, App
960	6	0.6	74	14	US-10-078-929-68	Sequence 68, Appl
961	6	0.6	75	9	US-09-864-761-45644	Sequence 45644, A
962	6	0.6	75	12	US-10-408-166-211	Sequence 211, App
963	6	0.6	76	9	US-09-864-761-44615	Sequence 44615, A
964	6	0.6	76	10	US-09-963-959-10	Sequence 10, Appl

```
 965   6  0.6   76  12  US-10-408-166-212      Sequence 212, App
 966   6  0.6   76  15  US-10-156-761-9828     Sequence 9828, Ap
 967   6  0.6   77  9   US-09-939-980-511      Sequence 511, App
 968   6  0.6   77  10  US-09-963-959-5        Sequence 5, Appli
 969   6  0.6   77  11  US-09-764-891-5129     Sequence 5129, Ap
 970   6  0.6   77  12  US-10-408-166-213      Sequence 213, App
 971   6  0.6   78  9   US-09-764-869-800      Sequence 800, App
 972   6  0.6   78  12  US-10-227-577-800      Sequence 800, App
 973   6  0.6   78  12  US-10-408-166-214      Sequence 214, App
 974   6  0.6   78  15  US-10-091-504-800      Sequence 800, App
 975   6  0.6   79  9   US-09-764-887-216      Sequence 216, App
 976   6  0.6   79  10  US-09-738-626-5321     Sequence 5321, Ap
 977   6  0.6   79  12  US-10-408-166-215      Sequence 215, App
 978   6  0.6   79  15  US-10-073-961-216      Sequence 216, App
 979   6  0.6   80  12  US-10-408-166-216      Sequence 216, App
 980   6  0.6   81  11  US-09-992-600A-38      Sequence 38, Appl
 981   6  0.6   81  11  US-09-924-340-38       Sequence 38, Appl
 982   6  0.6   81  12  US-09-992-095B-38      Sequence 38, Appl
 983   6  0.6   81  12  US-10-154-678-38       Sequence 38, Appl
 984   6  0.6   81  12  US-09-999-570-38       Sequence 38, Appl
 985   6  0.6   81  12  US-10-289-762-1167     Sequence 1167, Ap
 986   6  0.6   81  12  US-10-408-166-217      Sequence 217, App
 987   6  0.6   81  15  US-10-000-489-38       Sequence 38, Appl
 988   6  0.6   81  15  US-10-000-986-38       Sequence 38, Appl
 989   6  0.6   82  12  US-10-408-166-218      Sequence 218, App
 990   6  0.6   82  15  US-10-301-064-34       Sequence 34, Appl
 991   6  0.6   83  12  US-09-864-408A-5086    Sequence 5086, Ap
 992   6  0.6   83  12  US-10-408-166-219      Sequence 219, App
 993   6  0.6   84  12  US-09-833-245-1908     Sequence 1908, Ap
 994   6  0.6   84  12  US-09-833-245-1909     Sequence 1909, Ap
 995   6  0.6   84  12  US-10-408-166-220      Sequence 220, App
 996   6  0.6   85  9   US-09-925-297-462      Sequence 462, App
 997   6  0.6   85  12  US-10-029-386-30397    Sequence 30397, A
 998   6  0.6   85  12  US-10-408-166-221      Sequence 221, App
 999   6  0.6   86  12  US-10-029-386-31914    Sequence 31914, A
1000   6  0.6   86  12  US-10-316-194-48       Sequence 48, Appl
```

ALIGNMENTS

```
RESULT 1
US-09-884-696-2
; Sequence 2, Application US/09884696
; Publication No. US20030035809A1
; GENERAL INFORMATION:
;  APPLICANT: GEORGE, LISLE W
;  APPLICANT:  ANGELOS, JOHN A
;  APPLICANT:  HESS, JOHN F
;  TITLE OF INVENTION: MORAXELLA BOVIS CYTOTOXIN, CYTOTOXIN GENE, ANTIBODIES
;  TITLE OF INVENTION:  AND VACCINES FOR PREVENTION AND TREATMENT OF MORAXELLA
;  TITLE OF INVENTION:  BOVIS INFECTIONS
;  FILE REFERENCE: 481.06
;  CURRENT APPLICATION NUMBER: US/09/884,696
;  CURRENT FILING DATE:  2001-06-19
;  NUMBER OF SEQ ID NOS: 41
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
;   LENGTH: 927
;   TYPE: PRT
;   ORGANISM: Moraxella bovis
US-09-884-696-2

  Query Match            89.1%;  Score 826;  DB 11;  Length 927;
  Best Local Similarity  99.9%;  Pred. No. 0;
  Matches  926;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 MSNINVIKSNIQAGLNSTKSGLKNLYLAIPKDYDPQKGGTLNDFIKAADELGIARLAEEP 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MSNINVIKSNIQAGLNSTKSGLKNLYLAIPKDYDPQKGGTLNDFIKAADELGIARLAEEP 60

Qy         61 NHTETAKKSVDTVNQFLSLTQTGIAISATKLEKFLQKHSTNKLAKGLDSVENIDRKLGKA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 NHTETAKKSVDTVNQFLSLTQTGIAISATKLEKFLQKHSTNKLAKGLDSVENIDRKLGKA 120

Qy        121 SNVLSTLSSFLGTALAGIELDSLIKKGDAAPDALAKASIDLINEIIGNLSQSTQTIEAFS 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 SNVLSTLSSFLGTALAGIELDSLIKKGDAAPDALAKASIDLINEIIGNLSQSTQTIEAFS 180

Qy        181 SQLAKLGSTISQAKGFSNIGNKLQNLNFSKTNLGLEIITGLLSGISAGFALADKNASTGK 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 SQLAKLGSTISQAKGFSNIGNKLQNLNFSKTNLGLEIITGLLSGISAGFALADKNASTGK 240

Qy        241 KVAAGFELSNQVIGNVTKAISSYVLAQRVAAGLSTTGAVAALITSSIMLAISPLAFMNAA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 KVAAGFELSNQVIGNVTKAISSYVLAQRVAAGLSTTGAVAALITSSIMLAISPLAFMNAA 300

Qy        301 DKFNHANALDEFAKQFRKFGYDGDHLLAEYQRGVGTIEASLTTISTALGAVSAGVSAAAV 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 DKFNHANALDEFAKQFRKFGYDGDHLLAEYQRGVGTIEASLTTISTALGAVSAGVSAAAV 360

Qy        361 GSAVGTPIALLVAGVTGLISGILEASKQAMFESVANRLQGKILEWEKQNGGQNYFDKGYD 420
              ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GSAVGAPIALLVAGVTGLISGILEASKQAMFESVANRLQGKILEWEKQNGGQNYFDKGYD 420

Qy        421 SRYAAYLANNLKFLSELNKELEAERVIAITQQRWDNNIGELAGITKLGERIKSGKAYADA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 SRYAAYLANNLKFLSELNKELEAERVIAITQQRWDNNIGELAGITKLGERIKSGKAYADA 480

Qy        481 FEDGKKVEAGSNITLDAKTGIIDISNSNGKKTQALHFTSPLLTAGTESRERLTNGKYSYI 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 FEDGKKVEAGSNITLDAKTGIIDISNSNGKKTQALHFTSPLLTAGTESRERLTNGKYSYI 540

Qy        541 NKLKFGRVKNWQVTDGEASSKLDFSKVIQRVAETEGTDEIGLIVNAKAGNDDIFVGQGKM 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 NKLKFGRVKNWQVTDGEASSKLDFSKVIQRVAETEGTDEIGLIVNAKAGNDDIFVGQGKM 600

Qy        601 NIDGGDGHDRVFYSKDGGFGNITVDGTSATEAGSYTVNRKVARGDIYHEVVKRQETKVGK 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 NIDGGDGHDRVFYSKDGGFGNITVDGTSATEAGSYTVNRKVARGDIYHEVVKRQETKVGK 660

Qy        661 RTETIQYRDYELRKVGYGYQSTDNLKSVEEVIGSQFNDVFKGSKFNDIFHSGEGDDLLDG 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 RTETIQYRDYELRKVGYGYQSTDNLKSVEEVIGSQFNDVFKGSKFNDIFHSGEGDDLLDG 720

Qy        721 GAGDDRLFGGKGNDRLSGDEGDDLLDGGSGDDVLNGGAGNDVYIFRKGDGNDTLYDGTGN 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GAGDDRLFGGKGNDRLSGDEGDDLLDGGSGDDVLNGGAGNDVYIFRKGDGNDTLYDGTGN 780

Qy        781 DKLAFADANISDIMIERTKEGIIVKRNDHSGSINIPRWYITSNLQNYQSNKTDHKIEQLI 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 DKLAFADANISDIMIERTKEGIIVKRNDHSGSINIPRWYITSNLQNYQSNKTDHKIEQLI 840

Qy        841 GKDGSYITSDQIDKILQDKKDGTVITSQELKKLADENKSQKLSASDIASSLNKLVGSMAL 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GKDGSYITSDQIDKILQDKKDGTVITSQELKKLADENKSQKLSASDIASSLNKLVGSMAL 900

Qy        901 FGTANSVSSNALQPITQPTQGILAPSV 927
              |||||||||||||||||||||||||||
Db        901 FGTANSVSSNALQPITQPTQGILAPSV 927


RESULT 2
US-09-884-696-13
; Sequence 13, Application US/09884696
; Publication No. US20030035809A1
; GENERAL INFORMATION:
;  APPLICANT: GEORGE, LISLE W
;  APPLICANT:  ANGELOS, JOHN A
;  APPLICANT:  HESS, JOHN F
;  TITLE OF INVENTION: MORAXELLA BOVIS CYTOTOXIN, CYTOTOXIN GENE, ANTIBODIES
;  TITLE OF INVENTION:  AND VACCINES FOR PREVENTION AND TREATMENT OF MORAXELLA
;  TITLE OF INVENTION:  BOVIS INFECTIONS
;  FILE REFERENCE: 481.06
;  CURRENT APPLICATION NUMBER: US/09/884,696
;  CURRENT FILING DATE:  2001-06-19
```

```
;  NUMBER OF SEQ ID NOS: 41
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
;   LENGTH: 14
;   TYPE: PRT
;   ORGANISM: Moraxella bovis
US-09-884-696-13

  Query Match             1.5%;  Score 14;  DB 11;  Length 14;
  Best Local Similarity   100.0%;  Pred. No. 5.7e-06;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        705 FNDIFHSGEGDDLL 718
              ||||||||||||||
Db          1 FNDIFHSGEGDDLL 14

RESULT 3
US-09-884-696-5
; Sequence 5, Application US/09884696
; Publication No. US20030035809A1
; GENERAL INFORMATION:
;  APPLICANT: GEORGE, LISLE W
;  APPLICANT:  ANGELOS, JOHN A
;  APPLICANT:  HESS, JOHN F
;  TITLE OF INVENTION: MORAXELLA BOVIS CYTOTOXIN, CYTOTOXIN GENE, ANTIBODIES
;  TITLE OF INVENTION:  AND VACCINES FOR PREVENTION AND TREATMENT OF MORAXELLA
;  TITLE OF INVENTION:  BOVIS INFECTIONS
;  FILE REFERENCE: 481.06
;  CURRENT APPLICATION NUMBER: US/09/884,696
;  CURRENT FILING DATE:  2001-06-19
;  NUMBER OF SEQ ID NOS: 41
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
;   LENGTH: 1023
;   TYPE: PRT
;   ORGANISM: Escherichia coli
US-09-884-696-5

  Query Match             1.5%;  Score 14;  DB 11;  Length 1023;
  Best Local Similarity   100.0%;  Pred. No. 0.00032;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        379 ISGILEASKQAMFE 392
              ||||||||||||||
Db        405 ISGILEASKQAMFE 418

RESULT 4
US-09-884-696-6
; Sequence 6, Application US/09884696
; Publication No. US20030035809A1
; GENERAL INFORMATION:
;  APPLICANT: GEORGE, LISLE W
;  APPLICANT:  ANGELOS, JOHN A
;  APPLICANT:  HESS, JOHN F
;  TITLE OF INVENTION: MORAXELLA BOVIS CYTOTOXIN, CYTOTOXIN GENE, ANTIBODIES
;  TITLE OF INVENTION:  AND VACCINES FOR PREVENTION AND TREATMENT OF MORAXELLA
;  TITLE OF INVENTION:  BOVIS INFECTIONS
;  FILE REFERENCE: 481.06
;  CURRENT APPLICATION NUMBER: US/09/884,696
;  CURRENT FILING DATE:  2001-06-19
;  NUMBER OF SEQ ID NOS: 41
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
;   LENGTH: 12
;   TYPE: PRT
;   ORGANISM: Moraxella bovis
US-09-884-696-6

  Query Match             1.3%;  Score 12;  DB 11;  Length 12;
  Best Local Similarity   100.0%;  Pred. No. 0.00043;
  Matches   12;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        433 FLSELNKELEAE 444
              ||||||||||||
Db          1 FLSELNKELEAE 12

RESULT 5
US-09-884-696-4
; Sequence 4, Application US/09884696
; Publication No. US20030035809A1
; GENERAL INFORMATION:
;  APPLICANT: GEORGE, LISLE W
;  APPLICANT:  ANGELOS, JOHN A
;  APPLICANT:  HESS, JOHN F
;  TITLE OF INVENTION: MORAXELLA BOVIS CYTOTOXIN, CYTOTOXIN GENE, ANTIBODIES
;  TITLE OF INVENTION:  AND VACCINES FOR PREVENTION AND TREATMENT OF MORAXELLA
;  TITLE OF INVENTION:  BOVIS INFECTIONS
;  FILE REFERENCE: 481.06
;  CURRENT APPLICATION NUMBER: US/09/884,696
;  CURRENT FILING DATE:  2001-06-19
;  NUMBER OF SEQ ID NOS: 41
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
;   LENGTH: 956
;   TYPE: PRT
;   ORGANISM: Actinobacillus pleuropneumoniae
US-09-884-696-4

  Query Match             1.2%;  Score 11;  DB 11;  Length 956;
  Best Local Similarity   100.0%;  Pred. No. 0.25;
  Matches   11;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        369 ALLVAGVTGLI 379
              |||||||||||
Db        392 ALLVAGVTGLI 402

RESULT 6
US-09-305-924-13
; Sequence 13, Application US/09305924A
; Publication No. US20030091579A1
; GENERAL INFORMATION:
;  APPLICANT: Jack G. Manns
;  APPLICANT: Stephen D. Acres
;  APPLICANT: Richard Harland
;  TITLE OF INVENTION: METHODS OF RAISING ANIMALS FOR MEAT PRODUCTION
;  FILE REFERENCE: 9001-0048
;  CURRENT APPLICATION NUMBER: US/09/305,924A
;  CURRENT FILING DATE: 1999-05-05
;  EARLIER APPLICATION NUMBER: US 60/084,217
;  EARLIER FILING DATE: 1998-05-05
;  NUMBER OF SEQ ID NOS: 14
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 13
;   LENGTH: 695
;   TYPE: PRT
;   ORGANISM: GnRH
US-09-305-924-13

  Query Match             1.1%;  Score 10;  DB 11;  Length 695;
  Best Local Similarity   100.0%;  Pred. No. 1.7;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        265 LAQRVAAGLS 274
              ||||||||||
Db        359 LAQRVAAGLS 368

RESULT 7
US-09-884-696-3
; Sequence 3, Application US/09884696
; Publication No. US20030035809A1
; GENERAL INFORMATION:
```

```
; APPLICANT: GEORGE, LISLE W
; APPLICANT:  ANGELOS, JOHN A
; APPLICANT:  HESS, JOHN F
; TITLE OF INVENTION: MORAXELLA BOVIS CYTOTOXIN, CYTOTOXIN GENE, ANTIBODIES
; TITLE OF INVENTION:  AND VACCINES FOR PREVENTION AND TREATMENT OF MORAXELLA
; TITLE OF INVENTION:  BOVIS INFECTIONS
; FILE REFERENCE: 481.06
; CURRENT APPLICATION NUMBER: US/09/884,696
; CURRENT FILING DATE:  2001-06-19
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
;   LENGTH: 953
;   TYPE: PRT
;   ORGANISM: Pasteurella haemolytica
US-09-884-696-3

  Query Match             1.1%;  Score 10;  DB 11;  Length 953;
  Best Local Similarity   100.0%;  Pred. No. 2.3;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        265 LAQRVAAGLS 274
              ||||||||||
Db        283 LAQRVAAGLS 292


RESULT 8
US-10-369-493-20166
; Sequence 20166, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
;  APPLICANT: Cao, Yongwei
;  APPLICANT:  Hinkle, Gregory J.
;  APPLICANT:  Slater, Steven C.
;  APPLICANT:  Goldman, Barry S.
;  APPLICANT:  Chen, Xianfeng
;  TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
;  TITLE OF INVENTION:  PLANTS WITH IMPROVED PROPERTIES
;  FILE REFERENCE: 38-10(52052)B
;  CURRENT APPLICATION NUMBER: US/10/369,493
;  CURRENT FILING DATE:  2003-02-28
;  PRIOR APPLICATION NUMBER: US 60/360,039
;  PRIOR FILING DATE: 2002-02-21
;  NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 20166
;   LENGTH: 1017
;   TYPE: PRT
;   ORGANISM: No. US20030233675A1toc punctiforme
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: (1)..(1017)
;   OTHER INFORMATION: unsure at all Xaa locations
US-10-369-493-20166

  Query Match             1.1%;  Score 10;  DB 12;  Length 1017;
  Best Local Similarity   100.0%;  Pred. No. 2.5;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        718 LDGGAGDDRL 727
              ||||||||||
Db        389 LDGGAGDDRL 398


RESULT 9
US-10-156-761-12748
; Sequence 12748, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
;  APPLICANT: OMURA, SATOSHI
;  APPLICANT:  IKEDA, HARUO
;  APPLICANT:  ISHIKAWA, JUN
;  APPLICANT:  HORIKAWA, HIROSHI
;  APPLICANT:  SHIBA, TADAYOSHI
;  APPLICANT:  SAKAKI, YOSHIYUKI
;  APPLICANT:  HATTORI, MASAHIRA
;  TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
;  FILE REFERENCE: 249-262
;  CURRENT APPLICATION NUMBER: US/10/156,761
;  CURRENT FILING DATE:  2002-05-29
;  PRIOR APPLICATION NUMBER: JP 2001-204089
;  PRIOR FILING DATE: 2001-05-30
;  PRIOR APPLICATION NUMBER: JP 2001-272697
;  PRIOR FILING DATE: 2001-08-02
;  NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 12748
;   LENGTH: 259
;   TYPE: PRT
;   ORGANISM: Streptomyces avermitilis
US-10-156-761-12748

  Query Match             1.0%;  Score 9;  DB 15;  Length 259;
  Best Local Similarity   100.0%;  Pred. No. 6.4;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        276 TGAVAALIT 284
              |||||||||
Db        164 TGAVAALIT 172


RESULT 10
US-10-156-761-12788
; Sequence 12788, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
;  APPLICANT: OMURA, SATOSHI
;  APPLICANT:  IKEDA, HARUO
;  APPLICANT:  ISHIKAWA, JUN
;  APPLICANT:  HORIKAWA, HIROSHI
;  APPLICANT:  SHIBA, TADAYOSHI
;  APPLICANT:  SAKAKI, YOSHIYUKI
;  APPLICANT:  HATTORI, MASAHIRA
;  TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
;  FILE REFERENCE: 249-262
;  CURRENT APPLICATION NUMBER: US/10/156,761
;  CURRENT FILING DATE:  2002-05-29
;  PRIOR APPLICATION NUMBER: JP 2001-204089
;  PRIOR FILING DATE: 2001-05-30
;  PRIOR APPLICATION NUMBER: JP 2001-272697
;  PRIOR FILING DATE: 2001-08-02
;  NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 12788
;   LENGTH: 307
;   TYPE: PRT
;   ORGANISM: Streptomyces avermitilis
US-10-156-761-12788

  Query Match             0.9%;  Score 8;  DB 15;  Length 307;
  Best Local Similarity   100.0%;  Pred. No. 70;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        353 AGVSAAAV 360
              ||||||||
Db        109 AGVSAAAV 116


RESULT 11
US-10-156-761-14641
; Sequence 14641, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
;  APPLICANT: OMURA, SATOSHI
;  APPLICANT:  IKEDA, HARUO
;  APPLICANT:  ISHIKAWA, JUN
;  APPLICANT:  HORIKAWA, HIROSHI
;  APPLICANT:  SHIBA, TADAYOSHI
;  APPLICANT:  SAKAKI, YOSHIYUKI
```

```
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 14641
;  LENGTH: 408
;  TYPE: PRT
;  ORGANISM: Streptomyces avermitilis
US-10-156-761-14641

  Query Match             0.9%;  Score 8;  DB 15;  Length 408;
  Best Local Similarity   100.0%;  Pred. No. 92;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        357 AAAVGSAV 364
              ||||||||
Db        298 AAAVGSAV 305


RESULT 12
US-10-369-493-17978
; Sequence 17978, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT:  Hinkle, Gregory J.
; APPLICANT:  Slater, Steven C.
; APPLICANT:  Goldman, Barry S.
; APPLICANT:  Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION:  PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE:  2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 17978
;  LENGTH: 424
;  TYPE: PRT
;  ORGANISM: SPHINGOMONAS
;  FEATURE:
;  NAME/KEY: unsure
;  LOCATION: (1)..(424)
;  OTHER INFORMATION: unsure at all Xaa locations
US-10-369-493-17978

  Query Match             0.9%;  Score 8;  DB 12;  Length 424;
  Best Local Similarity   100.0%;  Pred. No. 95;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        149 AAPDALAK 156
              ||||||||
Db        247 AAPDALAK 254


RESULT 13
US-10-369-493-18633
; Sequence 18633, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT:  Hinkle, Gregory J.
; APPLICANT:  Slater, Steven C.
; APPLICANT:  Goldman, Barry S.
; APPLICANT:  Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION:  PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE:  2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 18633
;  LENGTH: 477
;  TYPE: PRT
;  ORGANISM: Halobacterium sp. NRC-1
US-10-369-493-18633

  Query Match             0.9%;  Score 8;  DB 12;  Length 477;
  Best Local Similarity   100.0%;  Pred. No. 1.1e+02;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        358 AAVGSAVG 365
              ||||||||
Db         15 AAVGSAVG 22


RESULT 14
US-10-369-493-16555
; Sequence 16555, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT:  Hinkle, Gregory J.
; APPLICANT:  Slater, Steven C.
; APPLICANT:  Goldman, Barry S.
; APPLICANT:  Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION:  PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE:  2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 16555
;  LENGTH: 489
;  TYPE: PRT
;  ORGANISM: Bacillus thuringiensis
;  FEATURE:
;  NAME/KEY: unsure
;  LOCATION: (1)..(489)
;  OTHER INFORMATION: unsure at all Xaa locations
US-10-369-493-16555

  Query Match             0.9%;  Score 8;  DB 12;  Length 489;
  Best Local Similarity   100.0%;  Pred. No. 1.1e+02;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        358 AAVGSAVG 365
              ||||||||
Db         15 AAVGSAVG 22


RESULT 15
US-10-369-493-17426
; Sequence 17426, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT:  Hinkle, Gregory J.
; APPLICANT:  Slater, Steven C.
; APPLICANT:  Goldman, Barry S.
; APPLICANT:  Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION:  PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
```

```
;  CURRENT FILING DATE:  2003-02-28
;  PRIOR APPLICATION NUMBER: US 60/360,039
;  PRIOR FILING DATE: 2002-02-21
;  NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 17426
;   LENGTH: 505
;   TYPE: PRT
;   ORGANISM: Bacillus halodurans
US-10-369-493-17426

  Query Match             0.9%;  Score 8;  DB 12;  Length 505;
  Best Local Similarity   100.0%;  Pred. No. 1.1e+02;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        358 AAVGSAVG 365
              ||||||||
Db         14 AAVGSAVG 21


RESULT 16
US-10-156-761-11340
; Sequence 11340, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
;  APPLICANT: OMURA, SATOSHI
;  APPLICANT:  IKEDA, HARUO
;  APPLICANT:  ISHIKAWA, JUN
;  APPLICANT:  HORIKAWA, HIROSHI
;  APPLICANT:  SHIBA, TADAYOSHI
;  APPLICANT:  SAKAKI, YOSHIYUKI
;  APPLICANT:  HATTORI, MASAHIRA
;  TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
;  FILE REFERENCE: 249-262
;  CURRENT APPLICATION NUMBER: US/10/156,761
;  CURRENT FILING DATE:  2002-05-29
;  PRIOR APPLICATION NUMBER: JP 2001-204089
;  PRIOR FILING DATE: 2001-05-30
;  PRIOR APPLICATION NUMBER: JP 2001-272697
;  PRIOR FILING DATE: 2001-08-02
;  NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 11340
;   LENGTH: 527
;   TYPE: PRT
;   ORGANISM: Streptomyces avermitilis
US-10-156-761-11340

  Query Match             0.9%;  Score 8;  DB 15;  Length 527;
  Best Local Similarity   100.0%;  Pred. No. 1.2e+02;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        347 ALGAVSAG 354
              ||||||||
Db         75 ALGAVSAG 82


RESULT 17
US-10-369-493-1471
; Sequence 1471, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
;  APPLICANT: Cao, Yongwei
;  APPLICANT:  Hinkle, Gregory J.
;  APPLICANT:  Slater, Steven C.
;  APPLICANT:  Goldman, Barry S.
;  APPLICANT:  Chen, Xianfeng
;  TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
;  TITLE OF INVENTION:  PLANTS WITH IMPROVED PROPERTIES
;  FILE REFERENCE: 38-10(52052)B
;  CURRENT APPLICATION NUMBER: US/10/369,493
;  CURRENT FILING DATE:  2003-02-28
;  PRIOR APPLICATION NUMBER: US 60/360,039
;  PRIOR FILING DATE: 2002-02-21
;  NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 1471
;   LENGTH: 773
;   TYPE: PRT
;   ORGANISM: Saccharomyces cerevisiae
US-10-369-493-1471

  Query Match             0.9%;  Score 8;  DB 12;  Length 773;
  Best Local Similarity   100.0%;  Pred. No. 1.7e+02;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         46 KAADELGI 53
              ||||||||
Db        667 KAADELGI 674


RESULT 18
US-10-369-493-10178
; Sequence 10178, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
;  APPLICANT: Cao, Yongwei
;  APPLICANT:  Hinkle, Gregory J.
;  APPLICANT:  Slater, Steven C.
;  APPLICANT:  Goldman, Barry S.
;  APPLICANT:  Chen, Xianfeng
;  TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
;  TITLE OF INVENTION:  PLANTS WITH IMPROVED PROPERTIES
;  FILE REFERENCE: 38-10(52052)B
;  CURRENT APPLICATION NUMBER: US/10/369,493
;  CURRENT FILING DATE:  2003-02-28
;  PRIOR APPLICATION NUMBER: US 60/360,039
;  PRIOR FILING DATE: 2002-02-21
;  NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 10178
;   LENGTH: 4327
;   TYPE: PRT
;   ORGANISM: magnetite-containing magnetic coccus
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: (1)..(4327)
;   OTHER INFORMATION: unsure at all Xaa locations
US-10-369-493-10178

  Query Match             0.9%;  Score 8;  DB 12;  Length 4327;
  Best Local Similarity   100.0%;  Pred. No. 8.5e+02;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        720 GGAGDDRL 727
              ||||||||
Db       3976 GGAGDDRL 3983


RESULT 19
US-10-414-692-80
; Sequence 80, Application US/10414692
; Publication No. US20030228607A1
; GENERAL INFORMATION:
;  APPLICANT: X-Ceptor Therapeutics, Inc.
;  TITLE OF INVENTION: Screening method and modulators having an improved therapeutic
;  TITLE OF INVENTION:  profile
;  FILE REFERENCE: 8012-002-US
;  CURRENT APPLICATION NUMBER: US/10/414,692
;  CURRENT FILING DATE:  2003-04-14
;  PRIOR APPLICATION NUMBER: 60/372,650
;  PRIOR FILING DATE: 2002-04-15
;  NUMBER OF SEQ ID NOS: 86
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 80
;   LENGTH: 20
;   TYPE: PRT
;   ORGANISM: Artificial
;   FEATURE:
;   OTHER INFORMATION: Synthetic co-activator peptide
```

```
US-10-414-692-80

  Query Match             0.8%;  Score 7;  DB 12;  Length 20;
  Best Local Similarity   100.0%;  Pred. No. 50;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        744 LLDGGSG 750
              |||||||
Db         12 LLDGGSG 18


RESULT 20
US-08-424-550B-519
; Sequence 519, Application US/08424550B
; Publication No. US20020119447A1
;  GENERAL INFORMATION:
;    APPLICANT:  JOHN N. SIMONS
;    APPLICANT:  TAMI J. PILOT-MATIAS
;    APPLICANT:  GEORGE J. DAWSON
;    APPLICANT:  GEORGE G. SCHLAUDER
;    APPLICANT:  SURESH M. DESAI
;    APPLICANT:  THOMAS P. LEARY
;    APPLICANT:  ANTHONY SCOTT MUERHOFF
;    APPLICANT:  JAMES C. ERKER
;    APPLICANT:  SHERI L. BUIJK
;    APPLICANT:  ISA K. MUSHAHWAR
;    TITLE OF INVENTION:  NON-A, NON-B. NON-C, NON-D, NON-E HEPATITIS
;    TITLE OF INVENTION:  REAGENTS AND METHODS FOR THEIR USE
;    NUMBER OF SEQUENCES:  716
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  ABBOTT LABORATORIES  D377/AP6D
;      STREET:  100 ABBOTT PARK ROAD
;      CITY:  ABBOTT PARK
;      STATE:  IL
;      COUNTRY:  USA
;      ZIP:  60064-3500
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/424,550B
;      FILING DATE:
;      CLASSIFICATION:  435435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  POREMBSKI, PRISCILLA E.
;      REGISTRATION NUMBER:  33,207
;      REFERENCE/DOCKET NUMBER:  5527.PC.01
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  708-937-6365
;      TELEFAX:  708-938-2623
;  INFORMATION FOR SEQ ID NO:  519:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  44 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-424-550B-519

  Query Match             0.8%;  Score 7;  DB 8;  Length 44;
  Best Local Similarity   100.0%;  Pred. No. 1.1e+02;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        898 MALFGTA 904
              |||||||
Db         38 MALFGTA 44


RESULT 21
US-10-264-049-3945
; Sequence 3945, Application US/10264049
; Publication No. US20040005579A1
; GENERAL INFORMATION:
;  APPLICANT: Birse et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;  FILE REFERENCE: PA133P1
;  CURRENT APPLICATION NUMBER: US/10/264,049
;  CURRENT FILING DATE:  2002-10-04
;  PRIOR APPLICATION NUMBER: PCT/US01/18569
;  PRIOR FILING DATE: 2001-06-07
;  PRIOR APPLICATION NUMBER: US 60/209,467
;  PRIOR FILING DATE: 2000-06-07
;  NUMBER OF SEQ ID NOS: 4360
;  SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 3945
;   LENGTH: 49
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-264-049-3945

  Query Match             0.8%;  Score 7;  DB 12;  Length 49;
  Best Local Similarity   100.0%;  Pred. No. 1.2e+02;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        139 ELDSLIK 145
              |||||||
Db         17 ELDSLIK 23


RESULT 22
US-10-185-050-78
; Sequence 78, Application US/10185050
; Publication No. US20030077577A1
;   GENERAL INFORMATION:
;        APPLICANT: Pirozzi, Gregorio
;                   Kay, Brian K.
;                   Fowlkes, Dana M.
;        TITLE OF INVENTION: IDENTIFICATION AND ISOLATION OF NOVEL
;                            POLYPEPTIDES HAVING WW DOMAINS AND METHODS OF USING SAME
;        NUMBER OF SEQUENCES: 233
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: PENNIE & EDMONDS LLP
;             STREET: 1155 Avenue of the Americas
;             CITY: New  York
;             STATE: New York
;             COUNTRY: USA
;             ZIP: 10036-2711
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/185,050
;             FILING DATE: 28-Jun-2002
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: US/08/826,516
;             FILING DATE: 03-Apr-1997
;        ATTORNEY/AGENT INFORMATION:
;             NAME: MISROCK, S. LESLIE
;             REGISTRATION NUMBER: 18,872
;             REFERENCE/DOCKET NUMBER: 1101-208-999
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (212) 790-9090
;             TELEFAX: (212) 896-8864/9741
;             TELEX: 66141 PENNIE
;    INFORMATION FOR SEQ ID NO: 78:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 52 amino acids
;             TYPE: amino acid
;             STRANDEDNESS: <Unknown>
;             TOPOLOGY: unknown
;        MOLECULE TYPE: peptide
;        SEQUENCE DESCRIPTION: SEQ ID NO: 78:
```

```
US-10-185-050-78

  Query Match             0.8%;  Score 7;  DB 15;  Length 52;
  Best Local Similarity   100.0%;  Pred. No. 1.2e+02;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        243 AAGFELS 249
              |||||||
Db         46 AAGFELS 52


RESULT 23
US-09-917-340-9
; Sequence 9, Application US/09917340
; Patent No. US20020090369A1
; GENERAL INFORMATION:
;  APPLICANT: Murphy, Christopher J.
;  APPLICANT:  McAnulty, Jonathan F.
;  APPLICANT:  Reid, Ted W.
;  TITLE OF INVENTION: Transplant Media
;  FILE REFERENCE: TPLANT-06468
;  CURRENT APPLICATION NUMBER: US/09/917,340
;  CURRENT FILING DATE:  2001-07-29
;  PRIOR APPLICATION NUMBER: 60/221,632
;  PRIOR FILING DATE: 2000-07-28
;  PRIOR APPLICATION NUMBER: 60/249,602
;  PRIOR FILING DATE: 2000-11-17
;  PRIOR APPLICATION NUMBER: 60/290,932
;  PRIOR FILING DATE: 2001-05-15
;  NUMBER OF SEQ ID NOS: 96
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
;   LENGTH: 63
;   TYPE: PRT
;   ORGANISM: Bombyx mori
US-09-917-340-9

  Query Match             0.8%;  Score 7;  DB 9;  Length 63;
  Best Local Similarity   100.0%;  Pred. No. 1.5e+02;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        347 ALGAVSA 353
              |||||||
Db         16 ALGAVSA 22


RESULT 24
US-09-739-907-83
; Sequence 83, Application US/09739907
; Patent No. US20010012889A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: 36 Human Secreted Proteins
;  FILE REFERENCE: PZ022P1
;  CURRENT APPLICATION NUMBER: US/09/739,907
;  CURRENT FILING DATE:  2000-12-20
;  PRIOR APPLICATION NUMBER: 09/348,457
;  PRIOR FILING DATE: 1999-07-07
;  PRIOR APPLICATION NUMBER: 60/070,567
;  PRIOR FILING DATE: 1998-01-07
;  PRIOR APPLICATION NUMBER: 60/070,692
;  PRIOR FILING DATE: 1998-01-07
;  PRIOR APPLICATION NUMBER: 60/070,704
;  PRIOR FILING DATE: 1998-01-07
;  PRIOR APPLICATION NUMBER: 60/070,658
;  PRIOR FILING DATE: 1998-01-07
;  NUMBER OF SEQ ID NOS: 196
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 83
;   LENGTH: 89
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (89)
;   OTHER INFORMATION: Xaa equals stop translation
US-09-739-907-83

  Query Match             0.8%;  Score 7;  DB 9;  Length 89;
  Best Local Similarity   100.0%;  Pred. No. 2.1e+02;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        357 AAAVGSA 363
              |||||||
Db         52 AAAVGSA 58


RESULT 25
US-09-938-671-83
; Sequence 83, Application US/09938671
; Publication No. US20040002066A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: 36 Human Secreted Proteins
;  FILE REFERENCE: PZ022P1
;  CURRENT APPLICATION NUMBER: US/09/938,671
;  CURRENT FILING DATE:  2001-08-27
;  PRIOR APPLICATION NUMBER: 09/348,457
;  PRIOR FILING DATE: 1999-07-07
;  PRIOR APPLICATION NUMBER: 60/070,567
;  PRIOR FILING DATE: 1998-01-07
;  PRIOR APPLICATION NUMBER: 60/070,692
;  PRIOR FILING DATE: 1998-01-07
;  PRIOR APPLICATION NUMBER: 60/070,704
;  PRIOR FILING DATE: 1998-01-07
;  PRIOR APPLICATION NUMBER: 60/070,658
;  PRIOR FILING DATE: 1998-01-07
;  NUMBER OF SEQ ID NOS: 196
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 83
;   LENGTH: 89
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (89)
;   OTHER INFORMATION: Xaa equals stop translation
US-09-938-671-83

  Query Match             0.8%;  Score 7;  DB 12;  Length 89;
  Best Local Similarity   100.0%;  Pred. No. 2.1e+02;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        357 AAAVGSA 363
              |||||||
Db         52 AAAVGSA 58


RESULT 26
US-09-864-408A-5506
; Sequence 5506, Application US/09864408A
; Publication No. US20040009474A1
; GENERAL INFORMATION:
;  APPLICANT: Leach, Martin D.
;  APPLICANT:  Shimkets, Richard A.
;  TITLE OF INVENTION: No. US20040009474A1el Human Polynucleotides and Polypeptides Enco
;  FILE REFERENCE: 21402-012
;  CURRENT APPLICATION NUMBER: US/09/864,408A
;  CURRENT FILING DATE:  2001-05-24
;  PRIOR APPLICATION NUMBER: 60/206,690
;  PRIOR FILING DATE: 2000-05-24
;  NUMBER OF SEQ ID NOS: 9068
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5506
;   LENGTH: 90
;   TYPE: PRT
```

```
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(1)
;   OTHER INFORMATION: Wherein Xaa may be any naturally occuring amino acid
US-09-864-408A-5506

  Query Match             0.8%;  Score 7;  DB 12;  Length 90;
  Best Local Similarity   100.0%;  Pred. No. 2.1e+02;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        445 RVIAITQ 451
              |||||||
Db         31 RVIAITQ 37


RESULT 27
US-10-108-260A-3335
; Sequence 3335, Application US/10108260A
; Publication No. US20040005560A1
; GENERAL INFORMATION:
;  APPLICANT: HELIX RESEARCH INSTITUTE
;  TITLE OF INVENTION: No. US20040005560A1el full length cDNA
;  FILE REFERENCE: H1-A0106
;  CURRENT APPLICATION NUMBER: US/10/108,260A
;  CURRENT FILING DATE:  2002-03-27
;  NUMBER OF SEQ ID NOS: 5458
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3335
;   LENGTH: 100
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-108-260A-3335

  Query Match             0.8%;  Score 7;  DB 12;  Length 100;
  Best Local Similarity   100.0%;  Pred. No. 2.3e+02;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        371 LVAGVTG 377
              |||||||
Db         64 LVAGVTG 70


RESULT 28
US-09-924-358-20
; Sequence 20, Application US/09924358
; Patent No. US20020107376A1
; GENERAL INFORMATION:
;  APPLICANT: Millennium Pharmaceuticals, Inc.
;  APPLICANT:  Meyers, Rachel
;  APPLICANT:  MacBeth, Kyle
;  TITLE OF INVENTION: 26199, 33530, 33949, 47148, 50226, AND
;  TITLE OF INVENTION:  58764,
;  TITLE OF INVENTION:  NOVEL HUMAN TRANSFERASE FAMILY MEMBERS AND USES THEREFOR
;  FILE REFERENCE: 38155-20034.00
;  CURRENT APPLICATION NUMBER: US/09/924,358
;  CURRENT FILING DATE:  2001-08-06
;  PRIOR APPLICATION NUMBER: US 60/229,300
;  PRIOR FILING DATE: 2000-09-01
;  NUMBER OF SEQ ID NOS: 52
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 20
;   LENGTH: 102
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Consensus amino acid
US-09-924-358-20

  Query Match             0.8%;  Score 7;  DB 10;  Length 102;
  Best Local Similarity   100.0%;  Pred. No. 2.3e+02;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        761 DVYIFRK 767
              |||||||
Db         92 DVYIFRK 98


RESULT 29
US-10-410-764-20
; Sequence 20, Application US/10410764
; Publication No. US20040005664A1
; GENERAL INFORMATION:
;  APPLICANT: Millennium Pharmaceuticals, Inc.
;  APPLICANT:  Meyers, Rachel E.
;  APPLICANT:  MacBeth, Kyle J.
;  APPLICANT:  Curtis, Rory A.J.
;  APPLICANT:  Rudolph-Owen, Laura A.
;  APPLICANT:  Weich, Nadine S.
;  APPLICANT:  Olandt, Peter J.
;  APPLICANT:  Tsai, Fong-Ying
;  APPLICANT:  Kapeller-Libermann, Rosana
;  APPLICANT:  Carroll, Joseph M.
;  TITLE OF INVENTION: 26199, 33530, 33949, 47148, 50226,
;  TITLE OF INVENTION:  58764, 62113, 32144, 32235, 23565, 13305, 14911, 86216,
;  TITLE OF INVENTION:  25206 AND 8843 MOLECULES AND USES THEREFOR
;  FILE REFERENCE: MPI03-052OMNIM
;  CURRENT APPLICATION NUMBER: US/10/410,764
;  CURRENT FILING DATE:  2003-04-10
;  PRIOR APPLICATION NUMBER: US 09/924,358
;  PRIOR FILING DATE: 2001-08-06
;  PRIOR APPLICATION NUMBER: US 60/229,300
;  PRIOR FILING DATE: 2000-09-01
;  PRIOR APPLICATION NUMBER: US 10/350,553
;  PRIOR FILING DATE: 2003-01-24
;  PRIOR APPLICATION NUMBER: US 60/351,572
;  PRIOR FILING DATE: 2002-01-24
;  PRIOR APPLICATION NUMBER: US 09/966,614
;  PRIOR FILING DATE: 2001-09-27
;  PRIOR APPLICATION NUMBER: US 60/238,054
;  PRIOR FILING DATE: 2000-10-05
;  PRIOR APPLICATION NUMBER: US 10/281,094
;  PRIOR FILING DATE: 2002-10-25
;  PRIOR APPLICATION NUMBER: US 60/347,815
;  PRIOR FILING DATE: 2001-10-29
;  PRIOR APPLICATION NUMBER: US 10/076,535
;  PRIOR FILING DATE: 2002-02-15
;  PRIOR APPLICATION NUMBER: US 60/269,440
;  PRIOR FILING DATE: 2001-02-16
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 136
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 20
;   LENGTH: 102
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Amino acid consensus sequence
US-10-410-764-20

  Query Match             0.8%;  Score 7;  DB 12;  Length 102;
  Best Local Similarity   100.0%;  Pred. No. 2.3e+02;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        761 DVYIFRK 767
              |||||||
Db         92 DVYIFRK 98


RESULT 30
US-09-815-242-11108
; Sequence 11108, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
;  APPLICANT: Haselbeck, Robert
;  APPLICANT:  Ohlsen, Kari L.
```

```
; APPLICANT:  Zyskind, Judith W.
; APPLICANT:  Wall, Daniel
; APPLICANT:  Trawick, John D.
; APPLICANT:  Carr, Grant J.
; APPLICANT:  Yamamoto, Robert T.
; APPLICANT:  Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION:  Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE:  2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 11108
;   LENGTH: 103
;   TYPE: PRT
;   ORGANISM: Haemophilus influenzae
US-09-815-242-11108

  Query Match             0.8%;  Score 7;  DB 9;  Length 103;
  Best Local Similarity   100.0%;  Pred. No. 2.4e+02;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        481 FEDGKKV 487
              |||||||
Db         87 FEDGKKV 93

RESULT 31
US-09-741-669-297
; Sequence 297, Application US/09741669
; Patent No. US20020022718A1
; GENERAL INFORMATION:
; APPLICANT:  Forsyth, R. Allyn
; APPLICANT:  Ohlsen, Kari L.
; APPLICANT:  Zyskind, Judith W.
; TITLE OF INVENTION:  Genes identified as required for
; TITLE OF INVENTION:  proliferation of E. coli
; FILE REFERENCE:  ELITRA.009A
; CURRENT APPLICATION NUMBER: US/09/741,669
; CURRENT FILING DATE:  2000-12-19
; PRIOR APPLICATION NUMBER:   US 60/173005
; PRIOR FILING DATE:   1999-12-23
; NUMBER OF SEQ ID NOS:  481
; SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 297
;   LENGTH: 104
;   TYPE: PRT
;   ORGANISM: Escherichia coli
US-09-741-669-297

  Query Match             0.8%;  Score 7;  DB 9;  Length 104;
  Best Local Similarity   100.0%;  Pred. No. 2.4e+02;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        481 FEDGKKV 487
              |||||||
Db         87 FEDGKKV 93

RESULT 32
US-09-912-020-321
; Sequence 321, Application US/09912020
; Patent No. US20020045592A1
; GENERAL INFORMATION:
; APPLICANT: Zyskind, Judith
; APPLICANT:  Ohlsen, Kari L.
; APPLICANT:  Trawick, John
; APPLICANT:  Forsyth, R. Allyn
; APPLICANT:  Froelich, Jamie M.
; APPLICANT:  Carr, Grant J.
; APPLICANT:  Yamamoto, Robert T.
; APPLICANT:  Xu, H. Howard
; TITLE OF INVENTION: GENES IDENTIFIED AS REQUIRED FOR PROLIFERATION IN
; TITLE OF INVENTION:  ESCHERICHIA COLI
; FILE REFERENCE: ELITRA.001DV1
; CURRENT APPLICATION NUMBER: US/09/912,020
; CURRENT FILING DATE:  2001-07-23
; PRIOR APPLICATION NUMBER: 09/492,709
; PRIOR FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: 60/117,405
; PRIOR FILING DATE: 1999-01-27
; NUMBER OF SEQ ID NOS: 485
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 321Z
;   LENGTH: 104
;   TYPE: PRT
;   ORGANISM: E. Coli
US-09-912-020-321

  Query Match             0.8%;  Score 7;  DB 9;  Length 104;
  Best Local Similarity   100.0%;  Pred. No. 2.4e+02;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        481 FEDGKKV 487
              |||||||
Db         87 FEDGKKV 93

RESULT 33
US-09-815-242-10348
; Sequence 10348, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT:  Ohlsen, Kari L.
; APPLICANT:  Zyskind, Judith W.
; APPLICANT:  Wall, Daniel
; APPLICANT:  Trawick, John D.
; APPLICANT:  Carr, Grant J.
; APPLICANT:  Yamamoto, Robert T.
; APPLICANT:  Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION:  Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE:  2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
```

```
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 10348
;   LENGTH: 104
;   TYPE: PRT
;   ORGANISM: Escherichia coli
US-09-815-242-10348

  Query Match              0.8%;  Score 7;  DB 9;  Length 104;
  Best Local Similarity   100.0%;  Pred. No. 2.4e+02;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        481 FEDGKKV 487
              |||||||
Db         87 FEDGKKV 93


RESULT 34
US-09-815-242-14107
; Sequence 14107, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
;  APPLICANT: Haselbeck, Robert
;  APPLICANT:  Ohlsen, Kari L.
;  APPLICANT:  Zyskind, Judith W.
;  APPLICANT:  Wall, Daniel
;  APPLICANT:  Trawick, John D.
;  APPLICANT:  Carr, Grant J.
;  APPLICANT:  Yamamoto, Robert T.
;  APPLICANT:  Xu, H. Howard
;  TITLE OF INVENTION: Identification of Essential Genes in
;  TITLE OF INVENTION:  Prokaryotes
;  FILE REFERENCE: ELITRA.011A
;  CURRENT APPLICATION NUMBER: US/09/815,242
;  CURRENT FILING DATE:  2001-03-21
;  PRIOR APPLICATION NUMBER: 60/191,078
;  PRIOR FILING DATE: 2000-03-21
;  PRIOR APPLICATION NUMBER: 60/206,848
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 60/207,727
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: 60/242,578
;  PRIOR FILING DATE: 2000-10-23
;  PRIOR APPLICATION NUMBER: 60/253,625
;  PRIOR FILING DATE: 2000-11-27
;  PRIOR APPLICATION NUMBER: 60/257,931
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: 60/269,308
;  PRIOR FILING DATE: 2001-02-16
;  NUMBER OF SEQ ID NOS: 14110
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 14107
;   LENGTH: 104
;   TYPE: PRT
;   ORGANISM: Salmonella typhi
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: (1)...(104)
;   OTHER INFORMATION: Xaa = Any Amino Acid
US-09-815-242-14107

  Query Match              0.8%;  Score 7;  DB 9;  Length 104;
  Best Local Similarity   100.0%;  Pred. No. 2.4e+02;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        481 FEDGKKV 487
              |||||||
Db         87 FEDGKKV 93


RESULT 35
US-10-287-274-323
; Sequence 323, Application US/10287274
; Publication No. US20030181408A1
; GENERAL INFORMATION:
;  APPLICANT: Forsyth, R. Allyn
;  APPLICANT:  Ohlsen, Kari
;  APPLICANT:  Zyskind, Judith
;  TITLE OF INVENTION: GENES ESSENTIAL FOR MICROBIAL PROLIFERATION AND ANTISENSE THERETO
;  FILE REFERENCE: ELITRA.008DV1
;  CURRENT APPLICATION NUMBER: US/10/287,274
;  CURRENT FILING DATE:  2002-10-31
;  PRIOR APPLICATION NUMBER: US 60/164415
;  PRIOR FILING DATE: 1999-11-09
;  PRIOR APPLICATION NUMBER: US 09/711164
;  PRIOR FILING DATE: 2000-11-09
;  NUMBER OF SEQ ID NOS: 469
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 323
;   LENGTH: 104
;   TYPE: PRT
;   ORGANISM: Escherichia coli
US-10-287-274-323

  Query Match              0.8%;  Score 7;  DB 12;  Length 104;
  Best Local Similarity   100.0%;  Pred. No. 2.4e+02;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        481 FEDGKKV 487
              |||||||
Db         87 FEDGKKV 93


RESULT 36
US-09-989-919-93
; Sequence 93, Application US/09989919
; Patent No. US20020164344A1
; GENERAL INFORMATION:
;  APPLICANT: Macina, Roberto
;  APPLICANT:  Recipon, Herve
;  APPLICANT:  Pluta, Jason
;  APPLICANT:  Ghosh, Malavika
;  APPLICANT:  Sun, Yongming
;  APPLICANT:  Liu, Chenghua
;  TITLE OF INVENTION: Compositions and Methods Relating to Colon Specific Genes and Prot
;  FILE REFERENCE: DEX-0289
;  CURRENT APPLICATION NUMBER: US/09/989,919
;  CURRENT FILING DATE:  2001-11-21
;  PRIOR APPLICATION NUMBER: 60/252,505
;  PRIOR FILING DATE: 2000-11-22
;  NUMBER OF SEQ ID NOS: 124
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 93
;   LENGTH: 109
;   TYPE: PRT
;   ORGANISM: Homo sapien
US-09-989-919-93

  Query Match              0.8%;  Score 7;  DB 10;  Length 109;
  Best Local Similarity   100.0%;  Pred. No. 2.5e+02;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        559 SSKLDFS 565
              |||||||
Db         25 SSKLDFS 31


RESULT 37
US-09-864-761-40104
; Sequence 40104, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  APPLICANT:  Chen, Wensheng
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
```

```
;  TITLE OF INVENTION:  GENE EXPRESSION ANALYSIS BY MICROARRAY
;  FILE REFERENCE: Aeomica-X-1
;  CURRENT APPLICATION NUMBER: US/09/864,761
;  CURRENT FILING DATE:  2001-05-23
;  PRIOR APPLICATION NUMBER: US 60/180,312
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 09/632,366
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 40104
;   LENGTH: 114
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AC004132.1
;   OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 3.5
;   OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3.9
;   OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL =   3
;   OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 4.3
;   OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 5.4
;   OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 4.4
;   OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 5.5
;   OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 3.8
;   OTHER INFORMATION: EST_HUMAN HIT: AU129622.1, EVALUE 5.00e-42
;   OTHER INFORMATION: SWISSPROT HIT: P35680, EVALUE 5.00e-43
US-09-864-761-40104

  Query Match             0.8%;  Score 7;  DB 9;  Length 114;
  Best Local Similarity   100.0%;  Pred. No. 2.6e+02;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        735 RLSGDEG 741
              |||||||
Db         73 RLSGDEG 79


RESULT 38
US-10-264-049-3972
; Sequence 3972, Application US/10264049
; Publication No. US20040005579A1
; GENERAL INFORMATION:
;  APPLICANT: Birse et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;  FILE REFERENCE: PA133P1
;  CURRENT APPLICATION NUMBER: US/10/264,049
;  CURRENT FILING DATE:  2002-10-04
;  PRIOR APPLICATION NUMBER: PCT/US01/18569
;  PRIOR FILING DATE: 2001-06-07
;  PRIOR APPLICATION NUMBER: US 60/209,467
;  PRIOR FILING DATE: 2000-06-07
;  NUMBER OF SEQ ID NOS: 4360
;  SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 3972
;   LENGTH: 115
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: MISC_FEATURE
;   LOCATION: (98)
;   OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
;   FEATURE:
;   NAME/KEY: MISC_FEATURE
;   LOCATION: (106)
;   OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
US-10-264-049-3972

  Query Match             0.8%;  Score 7;  DB 12;  Length 115;
  Best Local Similarity   100.0%;  Pred. No. 2.6e+02;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        368 IALLVAG 374
              |||||||
Db          3 IALLVAG 9


RESULT 39
US-10-094-749-2223
; Sequence 2223, Application US/10094749
; Publication No. US20030219741A1
; GENERAL INFORMATION:
;  APPLICANT: ISOGAI, TAKAO
;  APPLICANT:  SUGIYAMA, TOMOYASU
;  APPLICANT:  OTSUKI, TETSUJI
;  APPLICANT:  WAKAMATSU, AI
;  APPLICANT:  SATO, HIROYUKI
;  APPLICANT:  ISHII, SHIZUKO
;  APPLICANT:  YAMAMOTO, JUN-ICHI
;  APPLICANT:  ISONO, YUUKO
;  APPLICANT:  HIO, YURI
;  APPLICANT:  OTSUKA, KAORU
;  APPLICANT:  NAGAI, KEIICHI
;  APPLICANT:  IRIE, RYOTARO
;  APPLICANT:  TAMECHIKA, ICHIRO
;  APPLICANT:  SEKI, NAOHIKO
;  APPLICANT:  YOSHIKAWA, TSUTOMU
;  APPLICANT:  OTSUKA, MOTOYUKI
;  APPLICANT:  NAGAHARI, KENJI
;  APPLICANT:  MASUHO, YASUHIKO
;  TITLE OF INVENTION: NOVEL FULL-LENGTH cDNA
;  FILE REFERENCE: 084335/0160
;  CURRENT APPLICATION NUMBER: US/10/094,749
;  CURRENT FILING DATE:  2002-03-12
;  PRIOR APPLICATION NUMBER: 60/350,435
;  PRIOR FILING DATE: 2002-01-24
;  PRIOR APPLICATION NUMBER: JP 2001-328381
;  PRIOR FILING DATE: 2001-09-14
;  NUMBER OF SEQ ID NOS: 3381
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2223
;   LENGTH: 116
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-094-749-2223
```

```
  Query Match             0.8%;  Score 7;  DB 12;  Length 116;
  Best Local Similarity   100.0%;  Pred. No. 2.6e+02;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        357 AAAVGSA 363
              |||||||
Db         15 AAAVGSA 21


RESULT 40
US-09-739-907-182
; Sequence 182, Application US/09739907
; Patent No. US20010012889A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: 36 Human Secreted Proteins
;  FILE REFERENCE: PZ022P1
;  CURRENT APPLICATION NUMBER: US/09/739,907
;  CURRENT FILING DATE:  2000-12-20
;  PRIOR APPLICATION NUMBER: 09/348,457
;  PRIOR FILING DATE: 1999-07-07
;  PRIOR APPLICATION NUMBER: 60/070,567
;  PRIOR FILING DATE: 1998-01-07
;  PRIOR APPLICATION NUMBER: 60/070,692
;  PRIOR FILING DATE: 1998-01-07
;  PRIOR APPLICATION NUMBER: 60/070,704
;  PRIOR FILING DATE: 1998-01-07
;  PRIOR APPLICATION NUMBER: 60/070,658
;  PRIOR FILING DATE: 1998-01-07
;  NUMBER OF SEQ ID NOS: 196
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 182
;   LENGTH: 118
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-739-907-182

  Query Match             0.8%;  Score 7;  DB 9;  Length 118;
  Best Local Similarity   100.0%;  Pred. No. 2.7e+02;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        357 AAAVGSA 363
              |||||||
Db         82 AAAVGSA 88


Search completed: February 17, 2004, 10:26:52
Job time : 51 secs
```

```
Db      408 GFELANQVVGNITKAVSSYILAQRVAAGLSSTGPVAALIASTVSLAISPLAFAGIADKFN 467

Qy      305 HANALDEFAKQFRKFGYDGDHLLAEYQRGVGTIEASLTTISTALGAVSAGVSAAAVGSAV 364
Db      468 HAKSLESYAERFKKLGYDGDNLLAEYQRGTGTIDASVTAINTALAAIAGGVSAAAAGSVI 527

Qy      365 GTPIALLVAGVTGLISGILEASKQAMFESVANRLQGKILEWEKQNGGQNYFDKGYDSRYA 424
Db      528 ASPIALLVSGITGVISTILQYSKQAMFEHVANKIHNKIVEWEKNNHGKNYFENGYDARYL 587

Qy      425 AYLANNLKFLSELNKELEAERVIAITQQRWDNNIGELAGITKLGERIKSGKAYADAFEDG 484
Db      588 ANLQDNMKFLLNLNKELQAERVIAITQQQWDNNIGDLAGISRLGEKVLSGKAYVDAFEEG 647

Qy      485 KKVEAGSNITLDAKTGIIDISNSNGKKTQALHFTSPLLTAGTESRERLTNGKYSYINKLK 544
Db      648 KHIKADKLVQLDSANGIIDVSNSGKAKTQHILFRTPLLTPGTEHRERVQTGKYEYITKLN 707

Qy      545 FGRVKNWQVTDGEASSKLDFSKVIQRV-------AETEGTDEIGLIVNAKAGNDDIFVGQ 597
Db      708 INRVDSWKITDGAASSTFDLTNVVQRIGIELDNAGNVTKTKETKIIAKLGEGDDNVFVGS 767

Qy      598 GKMNIDGGDGHDRVFYSKDGGFGNITVDGTSATEAGSYTVNRKVARGDIYHEVVKRQETK 657
Db      768 GTTEIDGGEGYDRVHYSR-GNYGALTIDATKETEQGSYTVNRFVETGKALHEVTSTHTAL 826

Qy      658 VGKRTETIQYRDYELRKVGYGYQSTDNLKSVEEVIGSQFNDVFKGSKFNDIFHSGEGDDL 717
Db      827 VGNREEKIEYR-HSNNQHHAGYYTKDTLKAVEEIIGTSHNDIFKGSKFNDAFNGGDGVDT 885

Qy      718 LDGGAGDDRLFGGKGNDRLSGDEGDDLLDGGSGDDVLNGGAGNDVYIFRKGDGNDTLYDG 777
Db      886 IDGNDGNDRLFGGKGDDILDGGNGDDFIDGGKGNDLLHGGKGDDIFVHRKGDGNDIITDS 945

Qy      778 TGNDKLAFADANISDIMIERTKEGIIVKRNDHSGSINIPRWY----ITSNLQNYQSNKTD 833
Db      946 DGNDKLSFSDSNLKDLTFEKVKHNLVI-TNSKKEKVTIQNWFREADFAKEVPNYKATK-D 1003

Qy      834 HKIEQLIGKDGSYITSDQIDKILQDKKDGTVITSQELKKLADENKSQKLSASDIASSLNK 893
Db     1004 EKIEEIIGQNGERITSKQVDDLI--AKGNGKITQDELSKVVDNYELLKHS-KNVTNSLDK 1060

Qy      894 LVGSMALFGTANSVSSNALQPITQPTQGI 922
Db     1061 LISSVSAFTSSNDSRNVLVAPTSMLDQSL 1089
```

RESULT 4
AAR22103

```
ID   AAR22103 standard; Protein; 1098 AA.
XX
AC   AAR22103;
XX
DT   06-JUL-1992  (first entry)
XX
DE   Bovine IL-2 - LKT fusion protein.
XX
KW   Interleukin 2; leuko-toxin; vaccine; pneumonia; respiratory
KW   diseases.
XX
OS   Pasteurella haemolytica.
OS   Bos taurus.
XX
PN   WO9203558-A.
XX
PD   05-MAR-1992.
XX
PF   22-AUG-1991;    91WO-CA00299.
XX
PR   22-AUG-1990;    90US-0571301.
XX
PA   (POTT/) POTTER A.
XX
PI   Potter A,  Campos M,  Hughes HPA;
XX
DR   WPI; 1992-096901/12.
DR   N-PSDB; AAQ22771.
XX
PT   Interleukin 2-leuko-toxin gene fusion - encodes fusion protein
PT   useful as vaccine for animal pneumonia
XX
PS   Claim 20; Fig 3; 68pp; English.
XX
CC   The IL-2-LKT protein was encoded by a chimeric gene contg. the
CC   bovine IL-2 gene fused to the DNA encoding at least one epitope of
CC   leukotoxin from P. haemolytica.  IL-2-LKT was gel purified and
CC   ligated into the expression vector pGH433 lacI.  The resulting
CC   clone pAA356 (ATCC 68386) contd. the desired gene fusion under the
CC   control of the E. coli lac promoter.  The protein produced by the
CC   gene fusion is useful in a vaccine compsn. with a pharmaceutically
CC   acceptable vehicle, e.g. a  carrier homologous to a rotavirus VP6
CC   inner capsid protein.  The vaccine can be used for preventing or
CC   ameliorating respiratory diseases in animals e.g. shipping fever
CC   or pneumonia.
CC   See also AAR24124,5.
XX
SQ   Sequence   1098 AA;

  Query Match             50.2%;  Score 2332;  DB 13;  Length 1098;
  Best Local Similarity   50.1%;  Pred. No. 9.9e-137;
  Matches  465;  Conservative  174;  Mismatches  264;  Indels   26;  Gaps   13;

Qy        8 KSNIQAGLNSTKSGLKNLYLAIPKD--YDPQKGGTLNDFIKAADELGIARLAEEPNHTET 65
Db      173 QSLTQAG-SSLKTGAKKIILYIPQNYQYDTEQGNGLQDLVKAAEELGIEVQREERNNIAT 231

Qy       66 AKKSVDTVNQFLSLTQTGIAISATKLEKFLQKHSTNKLAKGLDSVENIDRKLGKASNVLS 125
Db      232 AQTSLGTIQTAIGLTERGIVLSAPQIDKLLQK---TKAGQALGSAESIVQNANKAKTVLS 288

Qy      126 TLSSFLGTALAGIELDSLIKKGDAAPDALAKASIDLINEIIGNLSQSTQTIEAFSSQLAK 185
Db      289 GIQSILGSVLAGMDLDEAL-QNNSNQHALAKAGLELTNSLIENIANSVKTLDEFGEQISQ 347

Qy      186 LGSTISQAKGFSNIGNKLQNL-NFSKTNLGLEIITGLLSGISAGFALADKNASTGKKVAA 244
Db      348 FGSKLQNIKGLGTLGDKLKNIGGLDKAGLGLDVISGLLSGATAALVLADKNASTAKKVGA 407

Qy      245 GFELSNQVIGNVTKAISSYVLAQRVAAGLSTTGAVAALITSSIMLAISPLAFMNAADKFN 304
Db      408 GFELANQVVGNITKAVSSYILAQRVAAGLSSTGPVAALIASTVSLAISPLAFAGIADKFN 467

Qy      305 HANALDEFAKQFRKFGYDGDHLLAEYQRGVGTIEASLTTISTALGAVSAGVSAAAVGSAV 364
Db      468 HAKSLESYAERFKKLGYDGDNLLAEYQRGTGTIDASVTAINTALAAIAGGVSAAAAGSVI 527

Qy      365 GTPIALLVAGVTGLISGILEASKQAMFESVANRLQGKILEWEKQNGGQNYFDKGYDSRYA 424
Db      528 ASPIALLVSGITGVISTILQYSKQAMFEHVANKIHNKIVEWEKNNHGKNYFENGYDARYL 587

Qy      425 AYLANNLKFLSELNKELEAERVIAITQQRWDNNIGELAGITKLGERIKSGKAYADAFEDG 484
Db      588 ANLQDNMKFLLNLNKELQAERVIAITQQQWDNNIGDLAGISRLGEKVLSGKAYVDAFEEG 647

Qy      485 KKVEAGSNITLDAKTGIIDISNSNGKKTQALHFTSPLLTAGTESRERLTNGKYSYINKLK 544
Db      648 KHIKADKLVQLDSANGIIDVSNSGKAKTQHILFRTPLLTPGTEHRERVQTGKYEYITKLN 707

Qy      545 FGRVKNWQVTDGEASSKLDFSKVIQRV-------AETEGTDEIGLIVNAKAGNDDIFVGQ 597
Db      708 INRVDSWKITDGAASSTFDLTNVVQRIGIELDNAGNVTKTKETKIIAKLGEGDDNVFVGS 767

Qy      598 GKMNIDGGDGHDRVFYSKDGGFGNITVDGTSATEAGSYTVNRKVARGDIYHEVVKRQETK 657
Db      768 GTTEIDGGEGYDRVHYSR-GNYGALTIDATKETEQGSYTVNRFVETGKALHEVTSTHTAL 826
```

THIS PAGE BLANK (USPTO)

```
Db    173 QSLTQAG-SSLKTGAKKIILYIPQNYQYDTEQGNGLQDLVKAAEELGIEVQREERNNIAT 231

Qy     66 AKKSVDTVNQFLSLTQTGIAISATKLEKFLQKHSTNKLAKGLDSVENIDRKLGKASNVLS 125
Db    232 AQTSLGTIQTAIGLTERGIVLSAPQIDKLLQK---TKAGQALGSAESIVQNANKAKTVLS 288

Qy    126 TLSSFLGTALAGIELDSLIKKGDAAPDALAKASIDLINEIIGNLSQSTQTIEAFSSQLAK 185
Db    289 GIQSILGSVLAGMDLDEAL-QNNSNQHALAKAGLELTNSLIENIANSVKTLDEFGEQISQ 347

Qy    186 LGSTISQAKGFSNIGNKLQNL-NFSKTNLGLEIITGLLSGISAGFALADKNASTGKKVAA 244
Db    348 FGSKLQNIKGLGTLGDKLKNIGGLDKAGLGLDVISGLLSGATAALVLADKNASTAKKVGA 407

Qy    245 GFELSNQVIGNVTKAISSYVLAQRVAAGLSTTGAVAALITSSIMLAISPLAFMNAADKFN 304
Db    408 GFELANQVVGNITKAVSSYILAQRVAAGLSSTGPVAALIASTVSLAISPLAFAGIADKFN 467

Qy    305 HANALDEFAKQFRKFGYDGDHLLAEYQRGVGTIEASLTTISTALGAVSAGVSAAAVGSAV 364
Db    468 HAKSLESYAERFKKLGYDGDNLLAEYQRGTGTIDASVTAINTALAAIAGGVSAAAAGSVI 527

Qy    365 GTPIALLVAGVTGLISGILEASKQAMFESVANRLQGKILEWEKQNGGQNYFDKGYDSRYA 424
Db    528 ASPIALLVSGITGVISTILQYSKQAMFEHVANKIHNKIVEWEKNNHGKNYFENGYDARYL 587

Qy    425 AYLANNLKFLSELNKELEAERVIAITQQRWDNNIGELAGITKLGERIKSGKAYADAFEDG 484
Db    588 ANLQDNMKFLLNLNKELQAERVIAITQQQWDNNIGDLAGISRLGEKVLSGKAYVDAFEEG 647

Qy    485 KKVEAGSNITLDAKTGIIDISNSNGKKTQALHFTSPLLTAGTESRERLTNGKYSYINKLK 544
Db    648 KHIKADKLVQLDSANGIIDVSNSGKAKTQHILFRTPLLTPGTEHRERVQTGKYEYITKLN 707

Qy    545 FGRVKNWQVTDGEASSKLDFSKVIQRV-------AETEGTDEIGLIVNAKAGNDDIFVGQ 597
Db    708 INRVDSWKITDGAASSTFDLTNVVQRIGIELDNAGNVTKTKETKIIAKLGEGDDNVFVGS 767

Qy    598 GKMNIDGGDGHDRVFYSKDGGFGNITVDGTSATEAGSYTVNRKVARGDIYHEVVKRQETK 657
Db    768 GTTEIDGGEGYDRVHYSR-GNYGALTIDATKETEQGSYTVNRFVETGKALHEVTSTHTAL 826

Qy    658 VGKRTETIQYRDYELRKVGYGYQSTDNLKSVEEVIGSQFNDVFKGSKFNDIFHSGEGDDL 717
Db    827 VGNREEKIEYR-HSNNQHHAGYYTKDTLKAVEEIIGTSHNDIFKGSKFNDAFNGGDGVDT 885

Qy    718 LDGGAGDDRLFGGKGNDRLSGDEGDDLLDGGSGDDVLNGGAGNDVYIFRKGDGNDTLYDG 777
Db    886 IDGNDGNDRLFGGKGDDILDGGNGDDFIDGGKGNDLLHGGKGDDIFVHRKGDGNDIITDS 945

Qy    778 TGNDKLAFADANISDIMIERTKEGIIVKRNDHSGSINIPRWY----ITSNLQNYQSNKTD 833
Db    946 DGNDKLSFSDSNLKDLTFEKVKHNLVI-TNSKKEKVTIQNWFREADFAKEVPNYKATK-D 1003

Qy    834 HKIEQLIGKDGSYITSDQIDKILQDKKDGTVITSQELKKLADENKSQKLSASDIASSLNK 893
Db   1004 EKIEEIIGQNGERITSKQVDDLI--AKGNGKITQDELSKVVDNYELLKHS-KNVTNSLDK 1060

Qy    894 LVGSMALFGTANSVSSNALQPITQPTQGI 922
Db   1061 LISSVSAFTSSNDSRNVLVAPTSMLDQSL 1089
```

RESULT 3
AAB21073

```
ID   AAB21073 standard; Protein; 1098 AA.
XX
AC   AAB21073;
XX
DT   19-DEC-2000  (first entry)
XX
DE   Bovine IL-2/Pasteurella haemolytica leukotoxin fusion protein.
XX
KW   Bovine IL-2; interleukin-2; leukotoxin; LKT; respiratory disease;
KW   pneumonia; shipping fever; cattle; livestock; anti-Pasteurella vaccine;
KW   immunogen.
XX
OS   Chimeric - Bos taurus.
OS   Chimeric - Pasteurella haemolytica.
XX
PN   US6096320-A.
XX
PD   01-AUG-2000.
XX
PF   20-OCT-1997;   97US-0954418.
XX
PR   20-DEC-1993;   93US-0170126.
PR   22-JUL-1996;   96US-0681479.
PR   22-AUG-1990;   90US-0571301.
PR   16-OCT-1991;   91US-0777715.
XX
PA   (UYSA-) UNIV SASKATCHEWAN.
PA   (CIBA ) CIBA GEIGY CANADA LTD.
XX
PI   Campos M,  Hughes HPA,  Potter A;
XX
DR   WPI; 2000-531543/48.
DR   N-PSDB; AAA72483.
XX
PT   Vaccine for stimulating immunity against pneumonia comprises chimeric
PT   protein comprising gamma-interferon and leukotoxin derived from
PT   Pasteurella haemolytica  -
XX
PS   Example 1; Column 31-38; 56pp; English.
XX
CC   The invention relates to a novel vaccine composition comprising an
CC   immunogenic chimeric protein that comprises gamma-interferon (gamma-IFN)
CC   or an active fragment thereof, linked to an epitope of a Pasteurella
CC   haemolytica leukotoxin (LKT). Pasteurella species, especially Pasteurella
CC   haemolytica, are responsible for respiratory diseases in a range of
CC   agricultural animals, most particularly cattle, but also sheep, pigs,
CC   horses and fowl. Shipping fever is the most economically important
CC   respiratory disease associated with Pasteurella species, affecting
CC   15-30% of exposed cattle and resulting in a 2-5% mortality rate in the
CC   exposed population. The vaccine composition of the invention is
CC   is useful for preventing or ameliorating respiratory diseases such as
CC   pneumonia, particularly shipping fever pneumonia, in livestock. The
CC   present sequence represents a fusion protein comprising bovine
CC   interleukin-2 (IL-2) and Pasteurella haemolytica leukotoxin, which may
CC   also be used as an anti-Pasteurella vaccine.
XX
SQ   Sequence   1098 AA;

  Query Match             50.2%;  Score 2334;  DB 21;  Length 1098;
  Best Local Similarity   50.1%;  Pred. No. 7.4e-137;
  Matches  465;  Conservative  175;  Mismatches  263;  Indels   26;  Gaps   13;

Qy      8 KSNIQAGLNSTKSGLKNLYLAIPKD--YDPQKGGTLNDFIKAADELGIARLAEEPNHTET 65
Db    173 QSLTQAG-SSLKTGAKKIILYIPQNYQYDTEQGNGLQDLVKAAEELGIEVQREERNNIAT 231

Qy     66 AKKSVDTVNQFLSLTQTGIAISATKLEKFLQKHSTNKLAKGLDSVENIDRKLGKASNVLS 125
Db    232 AQTSLGTIQTAIGLTERGIVLSAPQIDKLLQK---TKAGQALGSAESIVQNANKAKTVLS 288

Qy    126 TLSSFLGTALAGIELDSLIKKGDAAPDALAKASIDLINEIIGNLSQSTQTIEAFSSQLAK 185
Db    289 GIQSILGSVLAGMDLDEAL-QNNSNQHALAKAGLELTNSLIENIANSVKTLDEFGEQISQ 347

Qy    186 LGSTISQAKGFSNIGNKLQNL-NFSKTNLGLEIITGLLSGISAGFALADKNASTGKKVAA 244
Db    348 FGSKLQNIKGLGTLGDKLKNIGGLDKAGLGLDVISGLLSGATAALVLADKNASTAKKVGA 407

Qy    245 GFELSNQVIGNVTKAISSYVLAQRVAAGLSTTGAVAALITSSIMLAISPLAFMNAADKFN 304
```

THIS PAGE BLANK (USPTO)

Baskar, P.
10/069799
Text Search


```
     FILE 'HCAPLUS, MEDLINE, BIOSIS, EMBASE, WPIDS, CONFSCI, SCISEARCH,
     JICST-EPLUS, JAPIO' ENTERED AT 09:57:43 ON 18 FEB 2004
L1           24 S "FARN J"?/AU
L2          390 S "STRUGNELL R"?/AU
L3          186 S "TENNENT J"?/AU
L4            7 S L1 AND L2 AND L3
L5            8 S L1 AND (L2 OR L3)
L6            7 S L2 AND L3
L7           49 S (L1 OR L2 OR L3) AND BOVIS
L8           50 S L4 OR L5 OR L6 OR L7
L9           13 DUP REM L8 (37 DUPLICATES REMOVED)
```

- Author(s)

L9 ANSWER 1 OF 13 HCAPLUS COPYRIGHT 2004 ACS on STN DUPLICATE 1

ACCESSION NUMBER: 2001:168028 HCAPLUS
DOCUMENT NUMBER: 134:221433
TITLE: Vaccine antigens of Moraxella
INVENTOR(S): **Farn, Jacinta; Strugnell, Richard; Tennent, Jan**
PATENT ASSIGNEE(S): Commonwealth Scientific and Industrial Research Organisation, Australia; The University of Melbourne
SOURCE: PCT Int. Appl., 60 pp. CODEN: PIXXD2
DOCUMENT TYPE: Patent
LANGUAGE: English
FAMILY ACC. NUM. COUNT: 1
PATENT INFORMATION:

```
     PATENT NO.        KIND  DATE           APPLICATION NO.  DATE
     ---------------   ----  --------       ---------------  --------
     WO 2001016172      A1   20010308       WO 2000-AU1048   20000831
         W:  AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH,
             CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH,
             GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK,
             LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ,
             PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ,
             UA, UG, US, UZ, VN, YU, ZA, ZW, AM, AZ, BY, KG, KZ, MD, RU,
             TJ, TM
        RW:  GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW, AT, BE, CH,
             CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE,
             BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG
     EP 1210364         A1   20020605        EP 2000-955974   20000831
         R:  AT, BE, CH, DE, DK, ES, FR, GB, GR, IT, LI, LU, NL, SE, MC,
             PT, IE, SI, LT, LV, FI, RO, MK, CY, AL
     BR 2000013574      A    20020611        BR 2000-13574    20000831
PRIORITY APPLN. INFO.:                    AU 1999-2571     A  19990831
                                          WO 2000-AU1048   W  20000831
```

AB The present invention relates to antigens of Moraxella, in particular, Moraxella **bovis**, nucleic acid sequences encoding these antigens and formulations for use in raising an immune response against Moraxella.

REFERENCE COUNT: 5 THERE ARE 5 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L9 ANSWER 2 OF 13 HCAPLUS COPYRIGHT 2004 ACS on STN DUPLICATE 2

ACCESSION NUMBER: 2001:811555 HCAPLUS
DOCUMENT NUMBER: 136:66038

TITLE: Molecular characterization of a secreted enzyme with phospholipase B activity from Moraxella **bovis**
AUTHOR(S): **Farn, Jacinta L.; Strugnell, Richard A.**; Hoyne, Peter A.; Michalski, Wojtek P.; **Tennent, Jan M.**
CORPORATE SOURCE: CSIRO Livestock Industries, Geelong, 3220, Australia
SOURCE: Journal of Bacteriology (2001), 183(22), 6717-6720
CODEN: JOBAAY; ISSN: 0021-9193
PUBLISHER: American Society for Microbiology
DOCUMENT TYPE: Journal
LANGUAGE: English
AB A candidate for a vaccine against infectious bovine keratoconjunctivitis (IBK) has been cloned and characterized from Moraxella **bovis**. The plb gene encodes a protein of 616 amino acids (mol. mass of .apprx.65.8 kDa) that expresses phospholipase B activity. Amino acid sequence anal. revealed that PLB is a new member of the GDSL (Gly-Asp-Ser-Leu) family of lipolytic enzymes.
REFERENCE COUNT: 32 THERE ARE 32 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L9 ANSWER 3 OF 13 HCAPLUS COPYRIGHT 2004 ACS on STN DUPLICATE 3
ACCESSION NUMBER: 2000:359019 HCAPLUS
DOCUMENT NUMBER: 133:100763
TITLE: Characterization of hemolysin of Moraxella **bovis** using a hemolysis-neutralizing monoclonal antibody
AUTHOR(S): Billson, F. Mark; Harbour, Colin; Michalski, Wojtek P.; **Tennent, Jan M.**; Egerton, John R.; Hodgson, Jennifer L.
CORPORATE SOURCE: Department of Veterinary Clinical Sciences, University of Sydney, Camden, 2570, Australia
SOURCE: Infection and Immunity (2000), 68(6), 3469-3474
CODEN: INFIBR; ISSN: 0019-9567
PUBLISHER: American Society for Microbiology
DOCUMENT TYPE: Journal
LANGUAGE: English
AB A concentrated bacterial culture supernatant from the hemolytic Moraxella **bovis** strain UQV 148NF was used to immunize mice and generate monoclonal antibodies (MAbs). One, MAb G3/D7, neutralized the hemolytic activity of M. **bovis** and recognized a 94-kDa protein by Western blot anal. in hemolytic M. **bovis** strains representing each of the different fimbrial serogroups. Exposure of corneal epithelial cells to M. **bovis** concentrated culture supernatants demonstrated a role for an exotoxin in the pathogenesis of infectious bovine keratoconjunctivitis, while neutralization of hemolytic and cytotoxic activities by MAb G3/D7 implies that these activities are related or have common epitopes. The action of M. **bovis** hemolysin was further characterized in sheep erythrocyte prepns. with a binding step and Ca2+ required for lysis to proceed, similar to the RTX family of bacterial exotoxins. Neutralization of lytic activity in vitro is evidence for the presence of M. **bovis** antigens, which may be capable of protecting cattle from the development of infectious

bovine keratoconjunctivitis.
REFERENCE COUNT: 39 THERE ARE 39 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L9 ANSWER 4 OF 13 HCAPLUS COPYRIGHT 2004 ACS on STN DUPLICATE 4
ACCESSION NUMBER: 2000:507582 HCAPLUS
DOCUMENT NUMBER: 133:249416
TITLE: Identification of type 4 fimbriae in Actinobacillus pleuropneumoniae
AUTHOR(S): Zhang, Y.; **Tennent, J. M.**; Ingham, A.; Beddome, G.; Prideaux, C.; Michalski, W. P.
CORPORATE SOURCE: Private Bag 24, Australian Animal Health Laboratory, CSIRO Animal Health, Geelong, 3220, Australia
SOURCE: FEMS Microbiology Letters (2000), 189(1), 15-18
CODEN: FMLED7; ISSN: 0378-1097
PUBLISHER: Elsevier Science B.V.
DOCUMENT TYPE: Journal
LANGUAGE: English
AB Type 4 fimbriae have been identified on the cell surface of Actinobacillus pleuropneumoniae by electron microscopy and N-terminal sequencing anal. A. pleuropneumoniae type 4 fimbrial subunit protein, purified from cell cultures and from outer membrane prepns., reacted with polyclonal antibody raised against type 4 fimbriae of Moraxella **bovis** on Western blots. N-terminal sequence anal. of the purified 17 kDa type 4 fimbrial subunit protein, named ApfA, revealed the first 12 amino acids to be identical to those of other type 4 fimbrial subunit proteins.
REFERENCE COUNT: 23 THERE ARE 23 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L9 ANSWER 5 OF 13 HCAPLUS COPYRIGHT 2004 ACS on STN DUPLICATE 5
ACCESSION NUMBER: 1997:19161 HCAPLUS
DOCUMENT NUMBER: 126:57234
TITLE: Identification, purification, and characterization of the type 4 fimbriae of Pasteurella multocida
AUTHOR(S): Ruffolo, Carmel G.; **Tennent, Jan M.**; Michalski, Wojtek P.; Adler, Ben
CORPORATE SOURCE: Dep. Microbiology, Monash Univ., Clayton, 3168, Australia
SOURCE: Infection and Immunity (1997), 65(1), 339-343
CODEN: INFIBR; ISSN: 0019-9567
PUBLISHER: American Society for Microbiology
DOCUMENT TYPE: Journal
LANGUAGE: English
AB The presence of fimbriae on Pasteurella multocida has been reported, but there have been no prior studies aimed at conclusively characterizing these structures. We now report on the identification and characterization of type 4 fimbriae on serogroup A, B, and D strains of P. multocida. Under microaerophilic conditions P. multocida showed an increased expression of the fimbriae, which were observed to form bundles. Fimbriae purified by high-performance reverse-phase liquid chromatog. constituted a single 18-kDa subunit, the first 21 amino acids of which shared very high similarity with the N-terminal amino acids sequence of other type 4

fimbrial subunits. Antiserum against the P. multocida 18-kDa protein immunostained the type 4 fimbrial subunit of Moraxella **bovis** and Dichelobacter nodosus. Based on these observations we conclude that P. multocida possesses type 4 fimbriae and have designated the P. multocida fimbrial subunit Ptf/A.

L9 ANSWER 6 OF 13 BIOSIS COPYRIGHT 2004 BIOLOGICAL ABSTRACTS INC. on STN

ACCESSION NUMBER:	1996:399092 BIOSIS
DOCUMENT NUMBER:	PREV199699121448
TITLE:	Construction of a defined aromatic mutant of the gram-positive bacterium C. pseudotuberculosis.
AUTHOR(S):	Simmons, Cameron [Reprint author]; **Farn, Jacinta** [Reprint author]; **Strugnell, Richard** [Reprint author]; Hodgson, Adrian
CORPORATE SOURCE:	Dep. Microbiol., Univ. Melbourne, Parkville, VIC 3052, Australia
SOURCE:	Brown, F. [Editor]; Norrby, E. [Editor]; Burton, D. [Editor]; Mekalanos, J. [Editor]. Vaccines (Cold Spring Harbor), (1996) pp. 111-116. Vaccines (Cold Spring Harbor); Molecular approaches to the control of infectious diseases. Publisher: Cold Spring Harbor Laboratory Press, 10 Skyline Drive, Plainview, New York 11803, USA. Series: Vaccines (Cold Spring Harbor). Meeting Info.: Thirteenth Meeting. Cold Spring Harbor, New York, USA. September 13-17, 1995. ISSN: 0899-4056. ISBN: 0-87969-479-3.
DOCUMENT TYPE:	Book Conference; (Meeting) Book; (Book Chapter) Conference; (Meeting Paper)
LANGUAGE:	English
ENTRY DATE:	Entered STN: 3 Sep 1996 Last Updated on STN: 3 Sep 1996

L9 ANSWER 7 OF 13 MEDLINE on STN DUPLICATE 6

ACCESSION NUMBER:	95336598 MEDLINE
DOCUMENT NUMBER:	PubMed ID: 7542003
TITLE:	Antibody reactivity to mycobacterial 65 kDa heat shock protein: relevance to autoimmunity.
AUTHOR:	Karopoulos C; Rowley M J; Handley C J; **Strugnell R A**
CORPORATE SOURCE:	Department of Biochemistry, Monash University, Clayton, Victoria, Australia.
SOURCE:	Journal of autoimmunity, (1995 Apr) 8 (2) 235-48. Journal code: 8812164. ISSN: 0896-8411.
PUB. COUNTRY:	ENGLAND: United Kingdom
DOCUMENT TYPE:	Journal; Article; (JOURNAL ARTICLE)
LANGUAGE:	English
FILE SEGMENT:	Priority Journals
ENTRY MONTH:	199508
ENTRY DATE:	Entered STN: 19950905 Last Updated on STN: 19960129 Entered Medline: 19950818

AB Reactivity to the mycobacterial 65 kDa heat shock protein (HSP 65) has been implicated in the pathogenesis of adjuvant arthritis in the rat, and may be involved in the pathogenesis of rheumatoid arthritis

or other autoimmune diseases in humans. Accordingly this study sought quantitative or qualitative differences in the antibody reactivity to HSP 65 between normal controls, patients with the multisystem autoimmune diseases, rheumatoid arthritis (RA) and systemic lupus erythematosus (SLE) and patients with the mycobacterial infections, tuberculosis (TB) and leprosy. Levels of antibodies to recombinant HSP 65 in serum were measured by ELISA in normal subjects and in patients with RA, SLE, TB or leprosy. Antibody reactivity was examined by Western blotting using polypeptide fragments of HSP 65 derived by recombinant DNA techniques, or by digestion with trypsin or cyanogen bromide (CNBr). Reactivity to a synthetic peptide, the adjuvant arthritis T-cell epitope of HSP 65 (180-188), was tested by ELISA. High levels of antibodies to full length recombinant HSP 65 from Mycobacterium **bovis** were present in all the groups tested. By Western blot analysis, most reactivity with intact HSP 65 was retained in a 32 kDa tryptic fragment, judged by sequencing and size estimations to represent amino acid residues 118- approximately 388. This sequence included a major T-cell epitope for adjuvant arthritis (180-188), but these nine amino acids were not essential for B-cell reactivity since most sera also reacted with residues 188-540 which lack the T-cell epitope. Moreover, the 180-188 synthetic peptide was unreactive by ELISA, and did not inhibit reactivity with the intact recombinant HSP 65. In conclusion, most individuals had antibodies to mycobacterial HSP 65, presumably resulting from previous bacterial infections. The magnitude of the response was unrelated to the occurrence of systemic autoimmune disease, and the pattern of antibody reactivity with recombinant and proteolytic fragments of HSP 65 suggests that the major B-cell epitope is conformational and consists of discontinuous regions of the molecule.

L9 ANSWER 8 OF 13 MEDLINE on STN DUPLICATE 7

ACCESSION NUMBER: 96037961 MEDLINE
DOCUMENT NUMBER: 96037961 PubMed ID: 7571364
TITLE: The protective efficacy of cloned Moraxella **bovis** pili in monovalent and multivalent vaccine formulations against experimentally induced infectious bovine keratoconjunctivitis (IBK).
AUTHOR: Lepper A W; Atwell J L; Lehrbach P R; Schwartzkoff C L; Egerton J R; **Tennent J M**
CORPORATE SOURCE: CSIRO Division of Animal Health, Animal Health Research Laboratory, Parkville, Vic., Australia.
SOURCE: VETERINARY MICROBIOLOGY, (1995 Jul) 45 (2-3) 129-38. Journal code: 7705469. ISSN: 0378-1135.
PUB. COUNTRY: Netherlands
DOCUMENT TYPE: (CLINICAL TRIAL)
(CONTROLLED CLINICAL TRIAL)
Journal; Article; (JOURNAL ARTICLE)
LANGUAGE: English
FILE SEGMENT: Priority Journals
ENTRY MONTH: 199511
ENTRY DATE: Entered STN: 19951227
Last Updated on STN: 19951227
Entered Medline: 19951108

AB Calves were vaccinated with cloned Moraxella **bovis** pili of serogroup C (experiment 1) or B (experiment 2) either as a monovalent formulation or as part of a multivalent preparation with

pili of six other serogroups. Within 4 weeks of the second vaccine dose vaccinated calves and non-vaccinated controls were challenged via the ocular route with either virulent M. **bovis** strain Dal2d (serogroup C) or M. **bovis** strain 3WO7 (serogroup B) in experiments 1 and 2, respectively. Calves vaccinated with multivalent vaccines had significantly lower antibody titres than those vaccinated with monovalent preparations. Nevertheless, the levels of protection against infectious bovine keratoconjunctivitis (IBK) achieved with multivalent vaccines were 72% and 83% for the groups challenged with M. **bovis** strains of serogroups B and C, respectively. The serogroup C monovalent vaccine gave 100% protection against experimentally induced IBK and M. **bovis** isolates cultured from the eyes 6 days post-challenge were identified as belonging solely to serogroup C. Unexpectedly, only 25% protection was achieved against homologous strain challenge of calves that received the monovalent serogroup B vaccine. Furthermore, the majority of M. **bovis** isolates recovered from calves in this group belonged to serogroup C, as did half of those isolates cultured from the multivalent vaccinates. The remaining bacterial isolates from the latter group, together with all isolates from the non-vaccinated controls, belonged to serogroup B. Results are consistent with the hypothesis that derivatives of the serogroup B challenge inoculum had expressed serogroup C pilus antigen within 6 days of the challenge, possibly as a result of pilus gene inversion occurring in response to the presence of specific antibody in eye tissues and tears.

L9 ANSWER 9 OF 13 HCAPLUS COPYRIGHT 2004 ACS on STN DUPLICATE 8
ACCESSION NUMBER: 1994:694086 HCAPLUS
DOCUMENT NUMBER: 121:294086
TITLE: Characterization of pilin genes from seven serologically defined prototype strains of Moraxella **bovis**
AUTHOR(S): Atwell, John L.; **Tennent, Jan M.**; Lepper, Anthony W. D.; Elleman, Tom C.
CORPORATE SOURCE: Commonwealth Scientific and Industrial Research Organisation, Victoria, 3052, Australia
SOURCE: Journal of Bacteriology (1994), 176(16), 4875-82
CODEN: JOBAAY; ISSN: 0021-9193
DOCUMENT TYPE: Journal
LANGUAGE: English

AB Numerous field isolates of Moraxella **bovis** have previously been classified by serol. techniques into seven serogroups, each defined by homologous cross-reaction with antisera prepared against purified pili of a single prototype strain. The gene encoding pilin from each of the prototype strains has been characterized by nucleotide sequence determination The coding sequences show extensive homol. (70 to 80%) while the strains were also characterized. The presence of an addnl., partial pilin gene in each prototype strain was confirmed by Southern blot anal., and the partial pilin genes from two strains of one serogroup were characterized by sequence determination Features of the pilin gene sequences are considered in relation to pilin gene inversion and the serol. variants which may arise from inversion events.

L9 ANSWER 10 OF 13 HCAPLUS COPYRIGHT 2004 ACS on STN DUPLICATE 9
ACCESSION NUMBER: 1995:236681 HCAPLUS
TITLE: A haemolytic cell-free preparation of Moraxella

bovis confers protection against Infectious Bovine Keratoconjunctivitis

AUTHOR(S): Billson, F. Mark; Hodgson, Jennifer L.; Egerton, John R.; Lepper, Anthony W. D.; Michalski, Wojtek P.; Schwartzkoff, C. Leigh; Lehrbach, Philip R.; **Tennent, Jan M.**

CORPORATE SOURCE: University of Sydney, Department of Animal Health, Private Bag 3, Camden, NSW, 2570, Australia

SOURCE: FEMS Microbiology Letters (1994), 124(1), 69-74
CODEN: FMLED7; ISSN: 0378-1097

PUBLISHER: Elsevier

DOCUMENT TYPE: Journal

LANGUAGE: English

AB Protection conferred by a cell-free preparation from a haemolytic Moraxella **bovis** isolate, UQV 148NF, was compared to an equivalent fraction from a non-haemolytic M. **bovis** isolate, Gordon 26L3, and to a recombinant DNA-derived pili vaccine. Three groups of ten calves were vaccinated twice with one of the three prepns. and, together with ten non-vaccinated calves, challenged with virulent M. **bovis** isolate Dal 2d. Compared to the control group, significant protection was observed in the group receiving the pili vaccine and the group receiving the preparation from haemolytic isolate, UQV 148NF.

L9 ANSWER 11 OF 13 MEDLINE on STN DUPLICATE 10

ACCESSION NUMBER: 94055030 MEDLINE

DOCUMENT NUMBER: 94055030 PubMed ID: 7901935

TITLE: A Moraxella **bovis** pili vaccine produced by recombinant DNA technology for the prevention of infectious bovine keratoconjunctivitis.

AUTHOR: Lepper A W; Elleman T C; Hoyne P A; Lehrbach P R; Atwell J L; Schwartzkoff C L; Egerton J R; **Tennent J M**

CORPORATE SOURCE: CSIRO Division of Animal Health, Animal Health Research Laboratory, Parkville, Vic., Australia.

SOURCE: VETERINARY MICROBIOLOGY, (1993 Jul) 36 (1-2) 175-83.
Journal code: 7705469. ISSN: 0378-1135.

PUB. COUNTRY: Netherlands

DOCUMENT TYPE: Journal; Article; (JOURNAL ARTICLE)

LANGUAGE: English

FILE SEGMENT: Priority Journals

ENTRY MONTH: 199312

ENTRY DATE: Entered STN: 19940117
Last Updated on STN: 19970203
Entered Medline: 19931203

AB Pili (fimbriae) were prepared from Moraxella **bovis** strain Dalton 2d (Dal2d) and from a derivative of Pseudomonas aeruginosa K/2PfS that contained a plasmid-borne Dal2d pilin gene and produced pili having serogroup-specific identity to Dal2d. Nine calves were vaccinated with two doses each of 30 micrograms authentic M. **bovis** Dal2d pili in oil adjuvant and 10 calves were vaccinated with a similar dose of P. aeruginosa-derived Dal2d pili in the same formulation. All 19 calves and 10 non-vaccinated controls were challenged by instillation of 1 x 10(9) virulent M. **bovis** Dal2d cells into both conjunctival sacs 19 days after the second vaccine dose. The serological response to vaccination and the degree of protection against experimentally induced

infectious bovine keratoconjunctivitis (IBK) were assessed. None of the nine calves vaccinated with authentic M. **bovis** Dal2d pili developed IBK while two of those vaccinated with P. aeruginosa-derived Dal2d pili developed lesions which accounted for a mean group lesion score of 0.3. In contrast, 9 of the 10 non-vaccinated calves developed IBK lesions, the majority of which were progressive, required early treatment and accounted for a mean group lesion score of 1.5. These results demonstrate the potential of a relatively low dose of pili produced by recombinant DNA technology for development of an effective vaccine against IBK.

L9 ANSWER 12 OF 13 BIOSIS COPYRIGHT 2004 BIOLOGICAL ABSTRACTS INC. on STN

ACCESSION NUMBER:	1993:200425 BIOSIS
DOCUMENT NUMBER:	PREV199344096675
TITLE:	Purification of proteases by immobilised bacitracin A affinity chromatography.
AUTHOR(S):	Michalski, Wojtek P.; Crooks, Jennipher K.; Prowse, Stephen J.; **Tennent, Jan M.**; Lepper, Anthony W. D.
CORPORATE SOURCE:	CSIRO Div. Anim. Health, Anim. Health Res. Lab., Private Bag No. 1, Parkville, Victoria 3052, Australia
SOURCE:	Journal of Cellular Biochemistry Supplement, (1993) Vol. 0, No. 17 PART A, pp. 50. Meeting Info.: Keystone Symposium on Protein Purification and Biochemical Engineering. Santa Fe, New Mexico, USA. January 15-21, 1993. ISSN: 0733-1959.
DOCUMENT TYPE:	Conference; (Meeting)
LANGUAGE:	English
ENTRY DATE:	Entered STN: 16 Apr 1993 Last Updated on STN: 9 Jun 1993

L9 ANSWER 13 OF 13 MEDLINE on STN DUPLICATE 11

ACCESSION NUMBER:	93069932 MEDLINE
DOCUMENT NUMBER:	93069932 PubMed ID: 1359693
TITLE:	The protective efficacy of pili from different strains of Moraxella **bovis** within the same serogroup against infectious bovine keratoconjunctivitis.
AUTHOR:	Lepper A W; Moore L J; Atwell J L; **Tennent J M**
CORPORATE SOURCE:	CSIRO Division of Animal Health, Animal Health Research Laboratory, Parkville, Vic., Australia.
SOURCE:	VETERINARY MICROBIOLOGY, (1992 Sep) 32 (2) 177-87. Journal code: 7705469. ISSN: 0378-1135.
PUB. COUNTRY:	Netherlands
DOCUMENT TYPE:	Journal; Article; (JOURNAL ARTICLE)
LANGUAGE:	English
FILE SEGMENT:	Priority Journals
ENTRY MONTH:	199212
ENTRY DATE:	Entered STN: 19930122 Last Updated on STN: 19970203 Entered Medline: 19921218

AB Three groups of ten calves were each immunised with a total of 400 micrograms pili prepared from three separate strains of Moraxella **bovis** in Alhydrogel-oil adjuvant as two divided, equal doses

21 days apart. Groups 1 and 2 each received a monovalent vaccine made from strain 4L and S276R respectively, which belonged to pili serogroup A. Group 3 received vaccine made from pili of strain Maffl, belonging to serogroup F. A further group of ten calves served as non-vaccinated controls. Calves in groups 1 and 2 had developed serogroup A-specific antibody and those in group 3 developed serogroup F-specific antibody, and some evidence of cross-reacting antibody was also detected when measured by an agglutination test using formalin-killed piliated cells of serogroup A strain 4L. Although antibody titres measured against purified pili by ELISA were highest with homologous serogroup antigens, cross-reactive titres to shared epitopes of M. **bovis** pili were also detected by this method. Ocular challenge of the 40 calves with virulent M. **bovis** of serogroup A strain S276R was carried out 14 days after the second vaccine dose. All non-vaccinated calves developed infectious bovine keratoconjunctivitis (IBK). The percentage protection in groups 1 (strain 4L) and 2 (strain S276R) was 60% and 80% respectively (P less than 0.05), with mean lesion scores of 0.7 and 0.3 out of a possible 6.0. The percentage protection of calves in group 3 (strain Maffl) was only 30%, with a mean lesion score of 1.4 compared with 2.2 for non-vaccinated controls. The present findings, together with other evidence indicating that immunity to IBK is serogroup-specific, suggest that inclusion of pili from one representative strain from each of the seven Australian and British serogroups in a polyvalent, subunit vaccine should effectively protect the majority of cattle against IBK caused by most field strains of M. **bovis** encountered in Australia and the United Kingdom.

(FILE 'HCAPLUS, MEDLINE, BIOSIS, EMBASE, WPIDS, CONFSCI, SCISEARCH, JICST-EPLUS, JAPIO' ENTERED AT 10:01:50 ON 18 FEB 2004)

```
L10          45 S (L1 OR L2 OR L3) AND MORAXELLA
L11           0 S L10 NOT L8
```

FILE 'HOME' ENTERED AT 10:02:32 ON 18 FEB 2004